Industrial Enzymes

Industrial Enzymes

Structure, Function and Applications

Edited by

Julio Polaina
and
Andrew P. MacCabe

Instituto de Agroquímica y Tecnología de Alimentos, CSIC, Valencia, Spain

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-1-4020-5376-4 (HB)
ISBN 978-1-4020-5377-1 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Cover illustration: Crystal structure of xylanase B from *Bacillus* sp. BP-23.
Courtesy of Julia Sanz-Aparicio.

Printed on acid-free paper

CONTENTS

PREFACE

INDUSTRIAL ENZYMES IN THE 21st CENTURY

Man's use of enzymes dates back to the earliest times of civilization. Important human activities in primitive communities such as the production of certain types of foods and beverages, and the tanning of hides and skins to produce leather for garments, involved the application of enzyme activities, albeit unknowingly. However, not until the 19th century with the development of biochemistry and the pioneering work of a number of eminent scientists did the nature of enzymes and how they work begin to be clarified. In France Anselme Payen and Jean-François Persoz described the isolation of an amylolytic substance from germinating barley (1833). Shortly afterwards the Swedish chemist Jöns Jacob Berzelius coined the term catalysis (1835) to describe the property of certain substances to accelerate chemical reactions. In Germany the physiologist Theodor Schwann discovered the digestive enzyme pepsin (1836), Wilhelm Kühne proposed the term 'enzyme' (1877), and the brothers Hans and Eduard Buchner demonstrated that the transformation of glucose into ethanol could be carried out by chemical substances (enzymes) present in cell-free extracts of yeast (1897). In the 1870's the Danish chemist Christian Hansen succeeded in obtaining pure rennet from calves' stomachs, the use of which in cheese-making resulted in considerable improvements in both product quantity and quality. Shortly thereafter he industrialised the production of rennet thus setting in motion the first enzyme production industry.

During the 20th century the recognition that enzymes are proteins along with the design of techniques for their purification and analysis, principally the work of James B. Sumner and Kaj Linderstrøm-Lang, paved the way for the development of procedures for their industrial production and use. The nineteen-sixties witnessed two major breakthroughs that had a major impact on the enzyme industry: the commercialisation of glucoamylase which catalyses the production of glucose from starch with much greater efficiency than that of the chemical procedure of acid hydrolysis, and the launch of the first enzyme-containing detergents. The development of genetic engineering in the eighties provided the tools necessary for the production and commercialisation of new enzymes thus seeding a second explosive expansion to the current billion dollar enzyme industry. Recent advances in X-ray crystallography and other analytical methods in the field of protein chemistry along with the ever increasing amounts of biological information available from genomics

programs and molecular techniques such as directed evolution and gene and genome shuffling, are bringing powerful means to bear on the study and manipulation of enzyme structure and function. The search for improvements in existing enzyme-catalysed procedures, the need to develop new technologies and the increasing concern for responsible use and reuse of raw materials can be expected to stimulate not only the rational modification of enzymes to match specific requirements but also the design of new enzymes with totally novel properties.

The aim of this book is to provide in a single volume an updated revision of the most important types of industrial enzymes based on consideration of their physicochemical and catalytic properties, three-dimensional structure, and the range of current and foreseeable applications. The first section of this volume is dedicated to the carbohydrate active enzymes which are extensively used not only in many food industry applications (baking, beverage production, starch processing, *etc*.) but also in the industrial production of textiles, detergents, paper, ethanol, *etc*. The second section, on peptidases, begins with an introductory chapter about the MEROPS database which constitutes the current classification of reference for this important group of enzymes, and subsequent chapters review the most industrially relevant types of peptidases. The section on lipases places special emphasis on the increasing application of these enzymes in synthetic processes. Nucleic acid modifying activities are considered in the fourth section. Whilst the nature of the applications and scale of use of the latter are not yet comparable to those of the enzymes considered in the preceding sections, they are of growing in importance given the indispensability of some in highly specialised fields including basic and applied research, medicine, pharmaceuticals, agronomy and forensics. The final section considers a number of important enzymes that cannot be classified into any of the other sections.

We wish to thank everyone involved in making this book possible and hope that it will become a tool equally useful to researchers, industrialists and students.

Julio Polaina
Andrew P. MacCabe

CONTRIBUTORS

Adányi, Nóra, sect. E, ch. 25, p. 439
Adney, William S., sect. A, ch. 3, p. 35
Alberghina, Lilia, sect. C, ch. 16, p. 263
Alcalde, Miguel, sect. E, ch. 26, p. 459
Arrizubieta, María J., sect. E, ch. 32, p. 565
Ballesteros, Antonio, sect. A, ch. 9, p. 141
Banerjee, Anirban, sect. E, ch. 30, p. 529
Banerjee, Uttam Chand, sect. E, ch. 30, p. 529
Barna, Teréz, sect. E, ch. 25, p. 439
Barrett, Alan J., sect. B, ch. 10, p. 161
Brinch-Pedersen, Henrik, sect. E, ch. 29, p. 503
Bujnicki, Janusz M., sect. D, ch. 21, p. 355
Chao, Yun-Peng, sect. E, ch. 34, p. 597
Chern, Jong-Tzer, sect. E, ch. 34, p. 597
Chiang, Chung-Jen, sect. E, ch. 34, p. 597
Christakopoulos, Paul, sect. A, ch. 6, p. 83
Claveríe-Martín, Félix, sect. B, ch. 13, p. 207
Dey, Estera S., sect. C, ch. 20, p. 339
Díaz, Pilar, sect. A, ch. 5. p. 65
Ding, Shi-You, sect. A, ch, 3, p. 35
Divakar, Soundar, sect. C, ch. 17, p. 283
Donlon, John, sect. B, ch. 12, p. 197
Dridi, Walid, sect. D, ch. 22, p. 377
Drouin, Régen, sect. D, ch. 22, p. 377
Emri, Tamás, sect. E, ch. 25, p. 439
Francisco, José da Cruz, sect. C, ch. 20, p. 339
Gallardo, Óscar, sect. A, ch. 5, p. 65
Gómez de Segura, Aránzazu, sect. A, ch. 9, p. 141
Gotor, Vicente, sect. C, ch. 18, p. 301
Gotor-Fernández, Vicente, sect. C, ch. 18, p. 301
Gough, Simon P., sect. C, ch. 20, p. 339
Grzonka, Zbigniew, sect. B, ch. 11, p. 181
Gummadi, Sathyanarayana N., sect. A, ch. 7, p. 99
Himmel, Michael E., sect. A, ch. 3, p. 35

SECTION A

CARBOHYDRATE ACTIVE ENZYMES

CHAPTER 1

AMYLOLYTIC ENZYMES: TYPES, STRUCTURES AND SPECIFICITIES

MARTIN MACHOVIČ[1] AND ŠTEFAN JANEČEK[1,2*]

[1]*Institute of Molecular Biology, Slovak Academy of Sciences, Bratislava, Slovakia and*
[2]*Department of Biotechnologies, Faculty of Natural Sciences, University of St. Cyril and Methodius, Trnava, Slovakia*
[*]*stefan.janecek@savba.sk*

1. INTRODUCTION

Cellulose and starch are the most abundant polymers on Earth. They both consist of glucose monomer units which are, however, differently bound to form polymer chains: starch contains the glucose linked up by the α-glucosidic bonds, while the glucose in cellulose is bound by the β-glucosidic linkages. Therefore these two important sources of energy for animals, plants and micro-organisms are biochemically hydrolysed by two different groups of enzymes: starch by α-glycoside hydrolases, and cellulose by β-glycoside hydrolases. Starch (amylon in Greek) consists of two distinct fractions: amylose – linear α-1,4-linked glucans, and amylopectin – linear α-1,4-linked glucans branched with α-1,6 linkages (Ball *et al.*, 1996; Mouille *et al.*, 1996), therefore the enzymes responsible for its hydrolysis are called amylolytic enzymes or – simply – amylases. Amylolytic enzymes form a large group of enzymes among which the most common and best known are α-amylases, β-amylases and glucoamylases.

Since starch (like the structurally related glycogen) is an essential source of energy, amylolytic enzymes are produced by a great variety of living organisms (Vihinen and Mäntsälä, 1989). Although the different amylases mediate the same reaction – they all catalyse the cleavage of the α-glucosidic bonds in the same substrate – structurally and mechanistically they are quite different (MacGregor *et al.*, 2001). Both α-amylase and β-amylase adopt the structure of a TIM-barrel fold (for a review see Pujadas and Palau, 1999), *i.e.* their catalytic domain consists of a $(\beta/\alpha)_8$-barrel formed by 8 parallel β-strands surrounded by 8 α-helices

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 3–18

(Matsuura *et al*., 1984; Mikami *et al*., 1993). The barrels are, however, not similar in their details (Jespersen *et al*., 1991). Glucoamylase on the other hand possesses the structure of an $(\alpha/\alpha)_6$-barrel, consisting of an inner barrel composed of 6 α-helices which is surrounded by 6 more (Aleshin *et al*., 1992). Strands and helices of the $(\beta/\alpha)_8$-barrel domain as well as the helices of the $(\alpha/\alpha)_6$-barrel are connected by loop regions of various lengths.

Based on the similarities and differences in their primary structures, amylolytic enzymes have been classified into families of glycoside hydrolases (GH) (Henrissat, 1991): (i) α-amylases – family GH13; (ii) β-amylases – family GH14; and (iii) glucoamylases – family GH15. This classification, available on-line at the CAZy (Carbohydrate-Active enZymes) internet site (Coutinho and Henrissat, 1999), reflects the differences in the reaction mechanisms and catalytic machinery employed by the three types of amylase (Davies and Henrissat, 1995). Due to the enormous accumulation of new sequence data in recent years, α-amylase family GH13 has expanded so that it now contains almost 30 different enzymes and proteins (*e.g.* pullulanase, isoamylase, neopullulanase *etc*.) exhibiting sequence relatedness to α-amylases (MacGregor *et al*., 2001). At present all these enzymes are classified into families GH13, GH70 and GH77 which together constitute glycoside hydrolase clan GH-H (Coutinho and Henrissat, 1999). Moreover, families GH31 and GH57 contain a few amylolytic specificities with no sequence similarity to family GH13 (Henrissat and Bairoch, 1996).

The present review focuses on structural characteristics of the GH families of amylases. Its main goal is to provide a brief overview of the best-known glycoside hydrolases families GH13, GH14, GH15, GH31, GH57 GH70 and GH77. Emphasis is placed on the description of their: (i) specificities with regard to the EC numbers; (ii) three-dimensional structures; and (iii) catalytic domain architecture.

2. CLAN GH-H: FAMILIES GH13, GH70 AND GH77

A recent list of members of clan GH-H is shown in Table 1. There are not only hydrolases (EC 3) but also transferases and isomerases from enzyme classes 2 and 5, respectively (Fig. 1). The GH13, GH70 and GH77 families constitute the members of the GH-H clan – the so-called the α-amylase family (MacGregor *et al*., 2001). This clan now covers about 30 different enzyme specificities (MacGregor, 2005). All GH-H clan members share several characteristics: (i) the catalytic domain is formed by the $(\beta/\alpha)_8$-barrel fold (*i.e.* TIM-barrel) with a longer loop connecting strand β3 to helix α3 known as domain B; (ii) a common catalytic mechanism in which the β4-strand aspartate acts as a base (nucleophile) and the β5-strand glutamate acts as a proton donor (acid/base catalyst) with the help of the third residue, the β7-strand aspartate, essential for substrate binding (transition state stabiliser); (iii) they employ the retaining mechanism for the cleavage of the α-glycosidic bonds (Matsuura *et al.*, 1984; Buisson *et al*., 1987; Machius *et al*., 1995; Aghajari *et al*., 1998; Matsuura, 2002).

Besides the requirements for classification, it is practically impossible to study the α-amylase family without taking into account the conserved sequence regions

Table 1. α-Amylase family (clan GH-H)

Enzyma class	Enzyme	EC	GH family
Hydrolases	α-Amylase	3.2.1.1	13
	Oligo-1,6-glucosidase	3.2.1.10	13
	α-Glucosidase	3.2.1.20	13
	Pullulanase	3.2.1.41	13
	Amylopullulanase	3.2.1.1/41	13
	Cyclomaltodextrinase	3.2.1.54	13
	Maltotetraohydrolase	3.2.1.60	13
	Isoamylase	3.2.1.68	13
	Dextranglucosidase	3.2.1.70	13
	Trehalose-6-phosphate hydrolase	3.2.1.93	13
	Maltohexaohydrolase	3.2.1.98	13
	Maltotriohydrolase	3.2.1.116	13
	Maltogenic α-amylase	3.2.1.133	13
	Maltogenic amylase	3.2.1.133	13
	Neopullulanase	3.2.1.135	13
	Maltooligosyltrehalose hydrolase	3.2.1.141	13
	Maltopentaohydrolase	3.2.1.-	13
Transferases	Amylosucrase	2.4.1.4	13
	Glucosyltransferase	2.4.1.5	70
	Sucrosea phosphorylase	2.4.1.7	13
	Glucan branching enzyme	2.4.1.18	13
	Cyclodextrin glucanotransferase	2.4.1.19	13
	4-α-Glucanotransferase	2.4.1.25	13, 77
	Glucan debranching enzyme	2.4.1.25/3.2.1.33	13
	Alternansucrasee	2.4.1.140	70
	Maltosyltransferase	2.4.1.-	13
Isomerases	Isomaltulose synthase	5.4.99.11	13
	Trehalose synthase	5.4.99.15	13
	Maltooligosyltrehalose synthase	5.4.99.16	13

(Janeček, 2002). It has been known for some time that the sequence similarity is extremely low (about 10%) even for the α-amylases alone (*i.e.* for EC 3.2.1.1). This was described for α-amylases from different micro-organisms, plants, and animals (Nakajima *et al*., 1986). With subsequent expansion of the family, *i.e.* when many sequences from various sources and with different enzyme specificities became available, the number of identical residues among the α-amylase family enzymes had decreased to 8-10 amino acids by 1994 (Janeček, 1994; Svensson, 1994). The conserved sequence regions of those α-amylase family members whose three-dimensional structures have already been solved are shown in Fig. 2. The regions of the GH70 glucan-synthesising glucosyltransferase are based on the prediction study by MacGregor *et al*., (1996) and site-directed mutagenesis (Devulapalle *et al*., 1997) since no three-dimensional structure is currently available for a GH70 member. It is clear that the GH-H clan contains the invariant catalytic triad consisting of two aspartates (in strands β4 and β7) and one glutamate (in strand β5). The two functionally important histidines (in strands β3 and β7) – although strongly conserved and

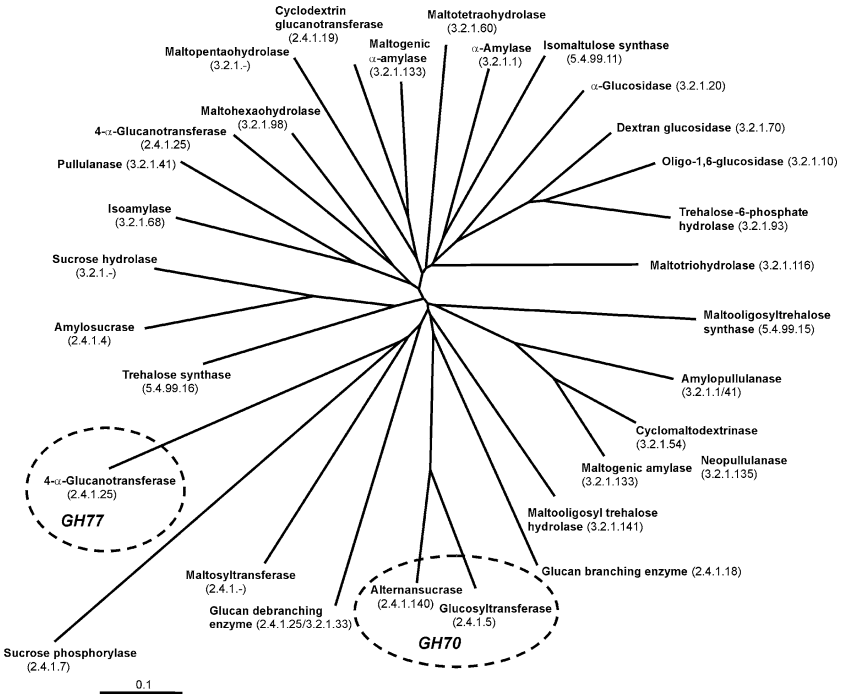


Figure 1. Evolutionary tree of the α-amylase family, *i.e.* clan GH-H. For the sake of simplicity, the tree is based on the alignment of conserved sequence regions (see Fig. 2), *i.e.* it does not reflect the complete amino acid sequences

apparently essential for several specificities (MacGregor *et al*., 2001) – are not present in GH13 maltosyltransferase (both His are missing) nor in the members of both GH70 and GH77 families (the β3 His is missing) (Fig. 2). The histidines have nevertheless been demonstrated to be critical in transition-state stabilisation (Søgaard *et al*., 1993). The fourth invariant residue of the α-amylase family seemed to be the arginine in the position $i-2$ with respect to the catalytic β4-strand aspartate (Janeček, 2002). However, this is no longer sustainable (Machovič and Janeček, 2003) because the sequences of the GH77 4-α-glucanotransferase from *Borrelia burgdorferi* and *Borrelia garinii* have the arginine substituted by a lysine (Fig. 3). This substitution is not a general feature characteristic of GH77 since it was not possible to detect more examples with such Arg/Lys substitution in the sequence databases. Moreover, the two putative *Borrelia* 4-α-glucanotransferases exhibit several additional remarkable sequence features that distinguish them from the rest of the GH77 enzymes. These are (Fig. 3): Pro/Ala in region VI (β2), Asp/Asn in region I (β3), Ile(Leu)/Trp and Leu-Gly/Phe-Gln(Glu) in region III (β5), and His/Gly in region IV (β7). With regard to protein function, catalytic activity and enzyme specificity of the two *Borrelia* 4-α-glucanotransferases, it

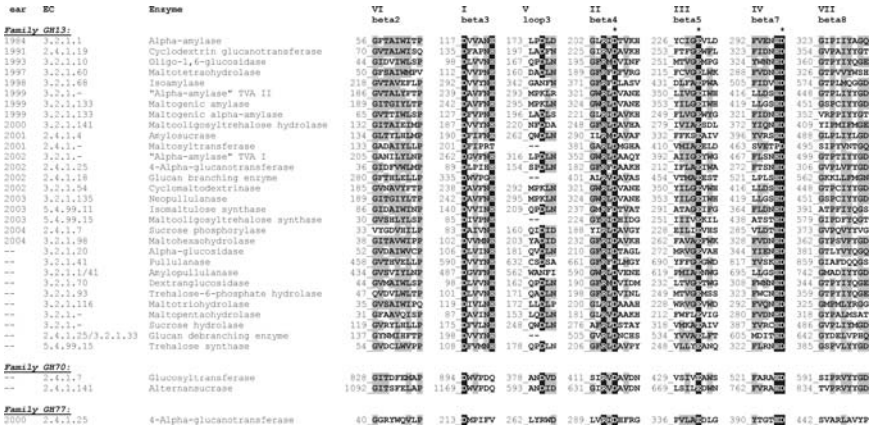


Figure 2. Conserved sequence regions in the α-amylase family. One representative for each enzyme specificity is presented. For those with three-dimensional structures already determined, the year when the structure was solved is shown in the first column. The important residues are highlighted in black; the catalytic triad is identified by asterisks. The other residues are coloured grey if conserved in at least 50% of the sequences. Figure adapted from Janeček (2002)

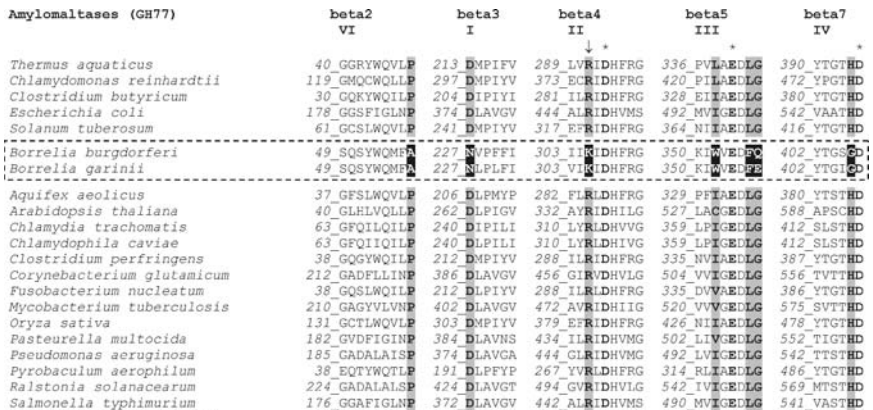


Figure 3. Selected conserved sequence regions in representative GH77 4-α-glucanotransferases. The regions I, II, III, IV and VI correspond to the strands β3, β4, β5, β7 and β2, respectively, of the catalytic $(\beta/\alpha)_8$-barrel domain. The members shown above the two *Borrelia* representatives are confirmed 4-α-glucanotransferases, whereas the members shown below are putative proteins only with GH77-like sequences. The invariant catalytic triad of the GH-H clan is identified by asterisks and bold characters. The important substituted residues in the two *Borrelia* 4-α-glucanotransferases are highlighted in black, the most interesting mutation (Arg/Lys) being emphasized by an arrow

is worth mentioning that these amino acid sequences were deduced from the nucleotide sequence of the Lyme disease spirochete and related genomes (Fraser *et al.*, 1997; Glöckner *et al.*, 2004), *i.e.* they are only translated ORFs. The 4-α-glucanotransferase specificities in both cases were thus assigned only by virtue of sequence similarities with other GH77 4-α-glucanotransferases/amylomaltases. The conserved catalytic triad, however, supports the possibility that the functions have been maintained. For example, the Arg/Lys mutant of *Bacillus stearothermophilus* α-amylase had 12% of the specific activity of the parental enzyme (Vihinen *et al.*, 1990) and the same mutant of the maize branching enzyme retained also some residual activity (Libessart and Preiss, 1998). The possibility of a sequencing error (Arg/Lys exchange) can be disregarded because the *Borrelia burgdorferi* 4-α-glucanotransferase was recently cloned, expressed in *Escherichia coli* and sequenced (Godany *et al.*, 2005). All the substitutions highlighted in Fig. 3 have been experimentally confirmed.

3. FAMILY GH13

GH13 ranks among the largest GH families with almost 30 enzyme specificities and more than 2,000 sequences (Coutinho and Henrissat, 1999; MacGregor *et al.*, 2001; Pujadas and Palau, 2001; Svensson *et al.*, 2002). It is the principal and most important family of the entire GH-H clan. In addition to α-amylase (EC 3.2.1.1), it contains (Table 1) cyclodextrin glucanotransferase (CGTase), α-glucosidase, amylopullulanase, neopullulanase, amylosucrase, *etc.* (MacGregor *et al.*, 2001). It seems reasonable to group the very closely related GH13 members into subfamilies, *e.g.* the oligo-1,6-glucosidase-like and neopullulanase-like members (Oslancova and Janeček, 2002; Oh, 2003).

Not all GH13 enzymes attack the glucosidic bonds in starch. However they do have a number of features in common (Svensson, 1994; Janeček, 1997; Kuriki and Imanaka, 1999; MacGregor, 2005): (i) sequence similarities (the so-called conserved sequence regions) covering the equivalent elements of their secondary structure (especially the β-strands); (ii) catalytic machinery (Asp, Glu and Asp residues in β-strands β4, β5 and β7, respectively); (iii) retaining reaction mechanism (the resulting hydroxyl group retains the α-configuration); (iv) the three-dimensional fold (TIM-barrel). The first three-dimensional structure of an α-amylase to be solved was that of Taka-amylase A, *i.e.* the α-amylase from *Aspergillus oryzae* (Matsuura *et al.*, 1984) (Fig. 4a). The enzyme adopts the so-called TIM-barrel fold which was first identified in the structure of triosephosphate isomerase (Banner *et al.*, 1975) and now found in about 50 different enzymes and proteins (Reardon and Farber, 1995; Janeček and Bateman, 1996; Pujadas and Palau, 1999). The $(\beta/\alpha)_8$-barrel motif consists of eight parallel β-strands forming the inner β-barrel which is surrounded by the outer cylinder composed of eight α-helices so that the individual β-strands and α-helices alternate and are connected by loops. Although all the members of the α-amylase family (Table 1) should share the characteristics given above, some have been classified into the new GH families (Coutinho and Henrissat, 1999). Thus

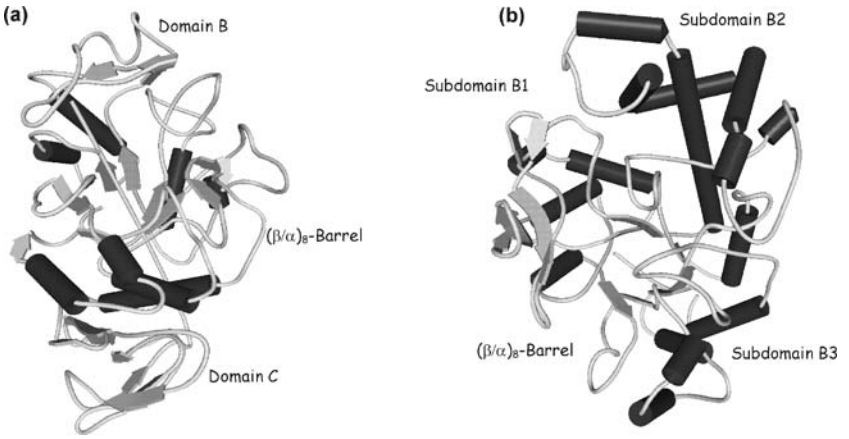


Figure 4. Three-dimensional structures of (***a***) GH13 α-amylase from *Aspergillus oryzae* (PDB code: 2TAA; Matsuura *et al*., 1984) and (***b***) GH77 amylomaltase from *Thermus aquaticus* (1CWY; Przylas *et al.*, 2000)

the sucrose-utilising glucosyltransferases (EC 2.4.1.5) have been placed in family GH70 because their catalytic domain was predicted to contain a circularly permuted version of the α-amylase type $(\beta/\alpha)_8$-barrel (MacGregor *et al*., 1996). This is also the case for one of the very recent members of the α-amylase family, alternansucrase (Argüello-Morales *et al*., 2000). Furthermore, some amylomaltases (EC 2.4.1.25), whose sequences exhibit low similarities with the most representative members of the α-amylase family, have been grouped into the new GH77 family (Coutinho and Henrissat, 1999). However, the three-dimensional structure of amylomaltase from *Thermus aquaticus* (Przylas *et al*., 2000) confirmed that this enzyme also possesses the regular $(\beta/\alpha)_8$-barrel structure (Fig. 4b) with the arrangement of the catalytic side-chains (two Asp residues and one Glu residue) being similar to that found in the α-amylase family.

With regard to quaternary structure, many members are able to form oligomers (Robyt, 2005). The most remarkable examples are cyclomaltodextrinases (for details, see Lee *et al*., 2005a; Turner *et al*., 2005).

4. FAMILIES GH14 AND GH15

There are two other amylolytic GH families in CAZy (Coutinho and Henrissat, 1999), GH14 and GH15, covering β-amylases and glucoamylases, respectively. They both employ the inverting mechanism for cleaving the α-glucosidic bonds, *i.e.* the products of their reactions are β-anomers (Sinnot, 1990; Kuriki, 2000; MacGregor *et al*., 2001). From an evolutionary point of view, β-amylases seem to be a 'solitary' GH family since they do not exhibit an obvious structural similarity to other glycoside hydrolases (Pujadas *et al*., 1996; Coutinho and Henrissat, 1999). By contrast, glucoamylases from GH15 form clan GH-L together with family GH65 (Egloff *et al*., 2001).

As regards sequence, these two types of amylase do not contain any of the conserved regions characteristic of the α-amylase family (Fig. 2). Although they are both exo-amylases their amino acid sequences and three-dimensional structures are different (Aleshin *et al*., 1992; Mikami *et al*., 1993). Structurally, β-amylase (Fig. 5a) ranks along with α-amylase among the large family of parallel $(\beta/\alpha)_8$-barrel proteins (Pujadas and Palau, 1999), while glucoamylase (Fig. 5b) belongs to a smaller family of proteins adopting the $(\alpha/\alpha)_6$-barrel fold (Aleshin *et al*., 1992).

Family GH14 includes β-amylases (EC 3.2.1.2) and hypothetical proteins with sequence similarity to β-amylases. Half of the family members are experimentally verified enzymes having β-amylase activity. β-Amylases are especially produced by plants: *Arabidopsis thaliana*, *Oryza sativa*, *Triticum aestivum* and *Solanum tuberosum*. Family GH15 includes glucoamylases (EC 3.2.1.3), two glucodextranases (EC 3.2.1.70) and hypothetical proteins with sequence similarity to GH15. Again, about 50% of the family members are experimentally verified enzymes having glucoamylase or glucodextranase activities.

The first determined three-dimensional structure of a β-amylase was that of soybean (Mikami *et al*., 1993). At present, the structures of β-amylases from sweet potato (Cheong *et al*., 1995), barley (Mikami *et al*., 1999b) and *Bacillus cereus* (Mikami *et al*., 1999a; Oyama *et al*., 1999) are also known. The core of the β-amylase structure is formed by the catalytic $(\beta/\alpha)_8$-barrel domain (Fig. 5a) followed by the C-terminal loop region. Although this loop surrounds the N-terminal side of the $(\beta/\alpha)_8$-barrel and may stabilise the whole β-amylase molecule, it is not involved in catalysis (Mikami, 2000). As has been pointed out above, the β-amylase $(\beta/\alpha)_8$-barrel differs from that of α-amylase and all other enzymes of clan GH-H, resembling more the single-domain structure of triosephosphate isomerase (Mikami, 2000). The two amino acid residues responsible for catalysis are the two glutamates, Glu186 and Glu380 (soybean β-amylase numbering), positioned

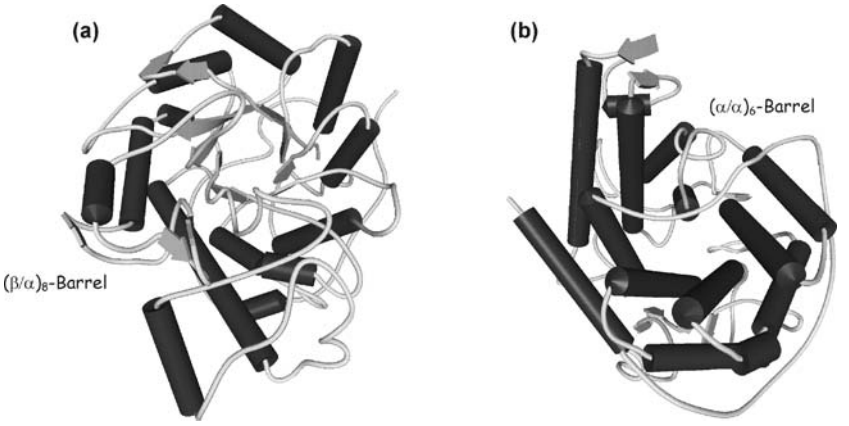


Figure 5. Three-dimensional structures of (***a***) GH14 β-amylase from soybean (1BYA; Mikami *et al.*, 1993) and (***b***) GH15 glucoamylase from *Aspergillus awamori* (1AGM; Aleshin *et al*., 1992)

near the C-terminus of strands β4 and β7 of the $(\beta/\alpha)_8$-barrel domain, respectively (Mikami *et al*., 1994). Totsuka and Fukazawa (1996) described further the indispensable roles for Asp101 and Leu383 in addition to the two catalytic glutamates. Analyses of the $(\beta/\alpha)_8$-barrel fold of β-amylases from both the evolutionary and structural points of view are available (Pujadas *et al*., 1996; Pujadas and Palau, 1997).

Glucoamylase structures have been solved for two fungal enzymes: *Aspergillus awamori* (Aleshin *et al*., 1992) and the yeast *Saccharomycopsis fibuligera* (Sevcik *et al*., 1998), and one bacterial enzyme from *Thermoanaerobacterium thermosaccharolyticum* (Aleshin *et al*., 2003). The glucoamylase catalytic domain is composed of 12 α-helices that form the so-called $(\alpha/\alpha)_6$-barrel fold (Fig. 5b). It consists of an inner core of six mutually parallel α-helices that are connected to each other through a peripheral set of six α-helices which are parallel to each other but approximately antiparallel to the inner core of the α-helices (Aleshin *et al*., 1992). This fold is not as frequent as the TIM-barrel fold (Farber and Petsko, 1990; Janeček and Bateman, 1996; Pujadas and Palau, 1999), however, the $(\alpha/\alpha)_6$-barrel has also been found in different proteins and enzymes, for example in the enzymes from families GH8 and GH9 (Juy *et al*., 1992; Alzari *et al*., 1996). Some glucoamylases, like some α-amylases (and related enzymes from the clan GH-H) and β-amylases, contain starch-binding domains (Svensson *et al*., 1989; Janeček and Sevcik, 1999) which can be of various types (for a review, see Rodriguez-Sanoja *et al*., 2005). The starch-binding domain may be evolutionarily independent from the catalytic domain (Janeček *et al*., 2003). It should also be possible to add a starch-binding domain artificially to an amylase (or eventually to any other protein) to improve its amylolytic and raw starch-binding and degradation abilities (Ohdan *et al*., 2000; Ji *et al*., 2003; Hua *et al*., 2004; Levy *et al*., 2004; Kramhøft *et al*., 2005; Latorre-Garcia *et al*., 2005). Recently, it seems evident that some amylases may contain starch-binding activity without a specific structural module (Hostinova *et al*., 2003; Tranier *et al*., 2005).

Based on the analysis of glucoamylase amino acid sequences, Coutinho and Reilly (1997) described seven subfamilies taxonomically corresponding to bacterial (1), archaeal (1), yeast (3) and fungal (2) origins. As evidenced by the crystal structures of the glucoamylases from *Aspergillus awamori* (Harris *et al*., 1993; Aleshin *et al*., 1994, 1996; Stoffer *et al*., 1995) and *Saccharomycopsis fibuligera* (Sevcik *et al*., 1998), the two glutamates, Glu179 and Glu400 (*Aspergillus* enzyme numbering), act as the key catalytic residues. The next most well-studied glucoamylase is that from *Aspergillus niger* (Christensen *et al*., 1996; Frandsen *et al*., 1996) which is highly similar to the *Aspergillus awamori* counterpart.

5. FAMILY GH31

There are some glucoamylases that have been classified into family GH31 together with α-glucosidases, α-xylosidases and glucan lyases (Yu *et al*., 1999; Lee *et al*., 2003; 2005b). These enzymes act through a retaining mechanism like the

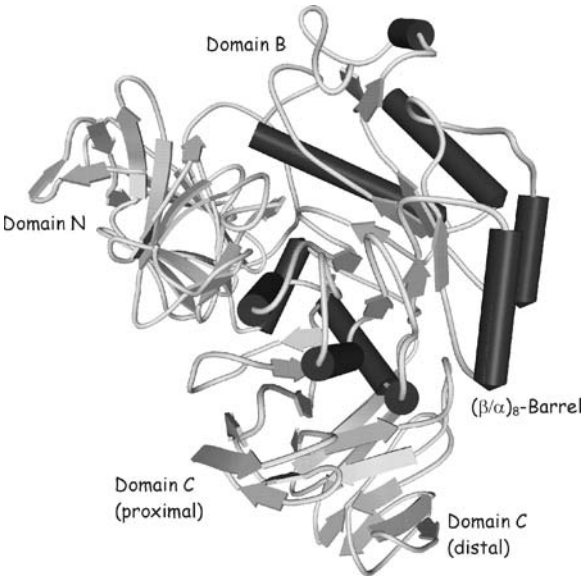


Figure 6. Three-dimensional structure of GH31 α-xylosidase from *Escherichia coli* (1XSI; Lovering *et al*, 2005)

members of clan GH-H (Chiba, 1997; Nakai *et al*, 2005). GH31 was considered to be a member of clan GH-H because of remote sequence homologies between GH31 and GH13 enzymes (Rigden, 2002). This assumption has recently been supported by the resolution of the three-dimensional structure of a GH31 α-xylosidase from *Escherichia coli* (Lovering *et al*, 2005) and α-glucosidase from *Sulfolobus solfataricus* (Ernst *et al*, 2006) showing the expected $(\beta/\alpha)_8$-barrel catalytic domain (Fig. 6). Interestingly, the domain arrangement of the GH31 members strongly resembles that of GH13 enzymes (Fig. 3), especially regarding domain B protruding out of the $(\beta/\alpha)_8$-barrel in the place of loop 3 (Lovering *et al*, 2005).

6. FAMILY GH57

For a long time GH57 has been one of the most popular GH families, attracting much scientific interest. More than 15 years ago the sequence of a heat-stable α-amylase from the thermophilic bacterium *Dictyoglomus thermophilum* was published (Fukusumi *et al*, 1988). Despite the fact that this sequence encoded an α-amylase, its analysis did not reveal any detectable similarity with GH13 α-amylases. Later, a similar sequence encoding the α-amylase from the hyperthermophilic archaeon, *Pyrococcus furiosus*, was determined (Laderman *et al*, 1993). These two sequences became the basis for the new amylolytic family, GH57, established in 1996 (Henrissat and Bairoch 1996). In the last few years, when entire genomes

of many micro-organisms have been sequenced, family GH57 has expanded. Its members are all prokaryotic enzymes, most of them from hyperthermophilic archaea (Zona *et al*., 2004). At present the GH57 family consists of about 100 members (Coutinho and Henrissat, 1999) and five enzyme (Janeček, 2005; Murakami *et al*., 2006): α-amylase (EC 3.2.1.1), α-galactosidase (EC 3.2.1.22), amylopullulanase (EC 3.2.1.1/41), branching enzyme (EC 2.4.1.18) and 4-α-glucanotransferase (EC 2.4.1.25). Only about 10% of the family sequence entries are enzymes; all others are hypothetical proteins without known activity (Zona *et al*., 2004). GH57 sequences are highly heterogeneous: some of them have less than 400 residues whereas others have more than 1,500 residues (Zona *et al*., 2004).

Structural information for GH57 members is scarce. To date, only the structures of the 4-α-glucanotransferase from *Thermococcus litoralis* (Imamura *et al*., 2003) and AmyC enzyme from *Thermotoga maritima* (Dickmanns *et al.*, 2006) have been determined. They both revealed a $(\beta/\alpha)_7$-barrel fold (Fig. 7), *i.e.* an incomplete TIM-barrel. Glu123 and Asp214 (*T. litoralis* enzyme numbering) which define the

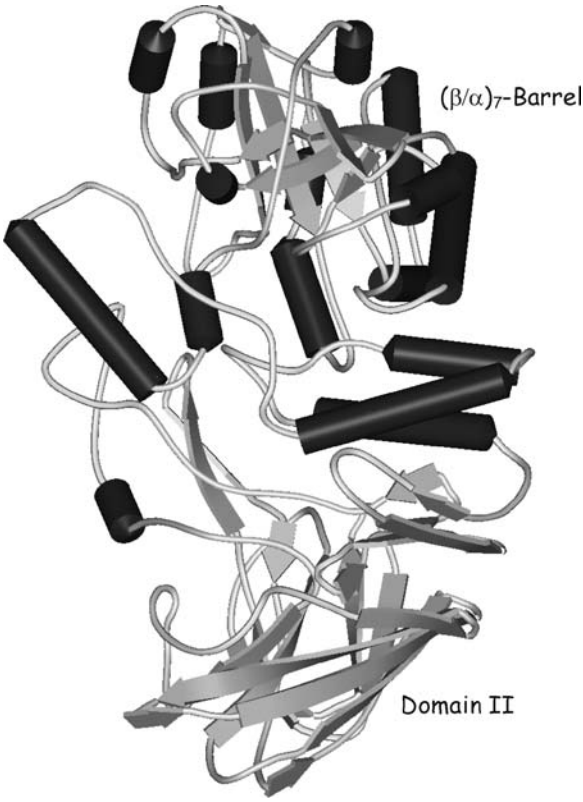


Figure 7. Three-dimensional structure of GH57 4-α-glucanotransferance from *Thermococcus litoralis* (1K1W; Imamura *et al*., 2003)

catalytic centre of the enzyme, are arranged at a distance of less than 7 Å (Imamura *et al.*, 2003), thus confirming that GH57 also employs a retaining mechanism for α-glycosidic bond cleavage.

New information about GH57 has arisen from a bioinformatic study focused on the conserved sequences containing the pair of catalytic residues (Zona *et al.*, 2004). In addition to *T. litoralis* 4-α-glucanotransferase, both catalytic residues were experimentally identified in two amylopullulanases from *Thermococcus hydrothermalis* (Zona *et al.*, 2004) and *Pyrococcus furiosus* (Kang *et al.*, 2005). The catalytic nucleophile was found also in the α-galactosidase from *Pyrococcus furiosus* (Van Lieshout *et al.*, 2003). Biochemical analysis indicates that family GH57 enzymes may lack a genuine α-amylase specificity (Janeček, 2005).

REFERENCES

Aghajari, N., Feller, G., Gerday, C., and Haser, R. (1998). Structures of the psychrophilic *Alteromonas haloplanctis* α-amylase give insights into cold adaptation at a molecular level. Structure *6*, 1503–1516.

Aleshin, A.E., Feng, P.H., Honzatko, R.B., and Reilly, P.J. (2003). Crystal structure and evolution of prokaryotic glucoamylase. J. Mol. Biol. *327*, 61–73.

Aleshin, A.E., Firsov, L.M., and Honzatko, R.B. (1994). Refined structure for the complex of acarbose with glucoamylase from *Aspergillus awamori* var. X100 to 2.4 Å resolution. J. Biol. Chem. *269*, 15631–15639.

Aleshin, A.E., Golubev, A., Firsov, L.M., and Honzatko, R.B. (1992). Crystal structure of glucoamylase from *Aspergillus awamori* var. X100 to 2.2 Å resolution. J. Biol. Chem. *267*, 19291–19298.

Aleshin, A.E., Stoffer, B., Firsov, L.M., Svensson, B., and Honzatko, R.B. (1996). Crystallographic complexes of glucoamylase with maltooligosaccharide analogs: relationships of stereochemical distorsions at the nonreducing end to the catalytic mechanism. Biochemistry *35*, 8319–8328.

Alzari, P.M., Souchon, H., and Dominguez, R. (1996). The crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from *Clostridium thermocellum*. Structure *4*, 265–275.

Argüello-Morales, M.A., Renaud-Simeon, M., Pizzut, S., Sarcabal, P., Willemot, R.M., and Monsan, P. (2000). Sequence analysis of the gene encoding alternansucrase from *Leuconostoc mesenteroides* NRRLB-1355. FEMS Microbiol. Lett. *182*, 81–85.

Ball, S., Guan, H.P., James, M., Myers, A., Keeling, P., Mouille, G., Buléon, A., Colonna, P., and Preiss, J. (1996). From glycogen to amylopectin: a model for the biogenesis of the plant starch granule. Cell *86*, 349–352.

Banner, D.W., Bloomer, A., Petsko, G.A., Phillips, D.C., and Wilson, I.A. (1975). Atomic coordinates for triose phosphate isomerase from chicken muscle. Biochem. Biophys. Res. Commun. *72*, 146–155.

Buisson, G., Duee, E., Haser, R., and Payan, F. (1987). Three dimensional structure of porcine pancreatic α-amylase at 2.9 Å resolution. Role of calcium in structure and activity. EMBO J. *6*, 3909–3916.

Cheong, C.G., Eom, S.H., Chang, C., Shin, D.H., Song, H.K., Min, K., Moon, J.H., Kim, K.K., Hwang, K.Y., and Suh, S. W. (1995). Crystallization, molecular replacement solution, and refinement of tetrameric β-amylase from sweet potato. Proteins *21*, 105–117.

Chiba, S. (1997). Molecular mechanism in α-glucosidase and glucoamylase. Biosci. Biotechnol. Biochem. *61*, 1233–1239.

Christensen, U., Olsen, K., Stoffer, B.B., and Svensson, B. (1996). Substrate binding mechanism of Glu180 → Gln, Asp176 → Asn, and wild-type glucoamylases from *Aspergillus niger*. Biochemistry *35*, 15009–15018.

Coutinho, P.M., and Henrissat, B. (1999). Glycoside hydrolase family server. URL: http://afmb.cnrs-mrs.fr/_pedro/CAZY/ghf.html

Coutinho, P.M., and Reilly, P.J. (1997). Glucoamylase structural, functional, and evolutionary relationships. Proteins *29*, 334–347.

Davies, G., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. Structure *3*, 853–859.

Devulapalle, K.S., Goodman, S.D., Gao, Q., Emsley, A. and Mooser, G. (1997). Knowledge-based model of a glucosyltransferase from the oral acterial group of mutans streptococci. Protein Sci. *6*, 2489–2493.

Dickmanns, A., Ballschmiter, M., Liebl, W., and Ficner, R. (2006). Structure of the novel α-amylase AmyC from *Thermotoga maritima*. Acta Crystallogr. D Biol. Crystallogr. *62*, 262–270.

Egloff, M.P., Uppenberg, J., Haalck, L., and Van Tilbeurgh, H. (2001). Crystal structure of maltose phosphorylase from *Lactobacillus brevis*: unexpected evolutionary relationship with glucoamylases. Structure *9*, 689–697.

Ernst, H.A., Lo Leggio, L., Willemoes, M., Leonard, G., Blum, P., and Larsen, S. (2006). Structure of the *Sulfolobus solfataricus* α-glucosidase: implications for domain conservation and substrate recognition in GH31. J. Mol. Biol. *358*, 1106–1124.

Farber, G.K., and Petsko G.A. (1990). The evolution of β/α barrel enzymes. Trends Biochem. Sci. *15*, 228–234.

Frandsen, T.P., Stoffer, B.B., Palcic, M.M., Hof, S., and Svensson, B. (1996). Structure and energetics of the glucoamylase-isomaltose transition-state complex probed by using modeling and deoxygenated substrates coupled with site-directed mutagenesis. J. Mol. Biol. *263*, 79–89.

Fraser, C.M., Casjens, S., and Venter, J.C. (1997). Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature *390*, 580–586.

Fukusumi, S., Kamizono, A., Horinouchi, S., and Beppu, T. (1988). Cloning and nucleotide sequence of a heat-stable amylase gene from an anaerobic thermophile, *Dictyoglomus thermophilum*. Eur. J. Biochem. *174*, 15–21.

Glöckner, G., Lehmann, R., Romualdi, A., Pradella, S., Schulte-Spechtel, U., Schilhabel, M., Wilske, B. Sühnel, J., and M. Platzer (2004). Comparative analysis of the *Borrelia garinii* genome. Nucleic Acids Res. *32*, 6038–6046.

Godany, A., Vidova, B., Bhide, M., Tkacikova, L., and Janeček, S. (2005). A unique member of the α-amylase family, the amylomaltase-like protein (BB0166) from *Borrelia burgdorferi* homologous to *Thermus aquaticus* amylomaltase, contains conserved only the catalytic triad. In: Thermophiles 05, International Conference, September 18-22, Gold Coast, Australia, pp. 149–150.

Harris, E. M. S., Aleshin, A.E., Firsov, L.M., and Honzatko, R.B. (1993). Refined structure for the complex of 1-deoxynojirimycin with glucoamylase from *Aspergillus awamori* var. X100 to 2.4 Å resolution. Biochemistry *32*, 1618–1626.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. *280*, 309–316.

Henrissat, B., and Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. Biochem. J. *316*, 695–696.

Hostinova, E., Solovicova, A., Dvorsky, R., and Gasperik, J. (2003). Molecular cloning and 3D structure prediction of the first raw-starch-degrading glucoamylase without a separate starch-binding domain. Arch. Biochem. Biophys. *411*, 189–195.

Hua, Y.W., Chi, M.C., Lo, H.F., Hsu, W.H., and Lin, L.L. (2004). Fusion of *Bacillus stearothermophilus* leucine aminopeptidase II with the raw-starch-binding domain of *Bacillus* sp. strain TS-23 α-amylase generates a chimeric enzyme with enhanced thermostability and catalytic activity. J. Ind. Microbiol. Biotechnol. *31*, 273–277.

Imamura, H., Fushinobu, S., Yamamoto, M., Kumasaka, T., Jeon, B.S., Wakagi, T., and Matsuzawa, H. (2003). Crystal structures of 4-α-glucanotransferase from *Thermococcus litoralis* and its complex with an inhibitor. J. Biol. Chem. *278*, 19378–19386.

Janeček, S. (1994). Parallel β/α-barrels of α-amylase, cyclodextrin glycosyltransferase and oligo-1,6-glucosidase versus the barrel of β-amylase: evolutionary distance is a reflection of unrelated sequences. FEBS Lett. *353*, 119–123.

Janeček, S. (1997). α-Amylase family: molecular biology and evolution. Progr. Biophys. Mol. Biol. *67*, 67–97.

Janeček, S. (2000). How many conserved sequence regions are there in the α-amylase family? Biologia, Bratislava *57 (Suppl. 11)*, 29–41.

Janeček, S. (2005). Amylolytic families of glycoside hydrolases: focus on the family GH-57. Biologia, Bratislava *60 (Suppl. 16)*, 177–184.

Janeček, S., and Bateman, A. (1996). The parallel $(\alpha/\beta)_8$-barrel: perhaps the most universal and the most puzzling protein folding motif. Biologia. Bratislava *51*, 613–628.

Janeček, S., and Sevcik, J. (1999). The evolution of starch-binding domain. FEBS Lett. *456*, 119–125.

Janeček, S., Svensson, B., and MacGregor, E. A. (2003). Relation between domain evolution, specificity, and taxonomy of the α-amylase family members containing a C-terminal starch-binding domain. Eur. J. Biochem. *270*, 635–645.

Jespersen, H., MacGregor, E.A., Sierks, M.R., and Svensson, B. (1991). Comparison of the domain-level organization of starch hydrolases and related enzymes. Biochem. J. *280*, 51–55.

Ji, Q., Vincken, J.P., Suurs, L.C., and Visser, R.G. (2003). Microbial starch-binding domains as a tool for targeting proteins to granules during starch biosynthesis. Plant Mol. Biol. *51*, 789–801.

Juy, M., Amit, A.G., Alzari, P.M., Poljak, R.J., Clayssens, M., Béguin, P., and Aubert J.P. (1992). Three-dimensional structure of a thermostable bacterial cellulase. Nature *357*, 89–91.

Kang, S., Vieille, C., and Zeikus, J.G. (2005). Identification of *Pyrococcus furiosus* amylopullulanase catalytic residues. Appl. Microbiol. Biotechnol. *66*, 408–413.

Kramhøft, B., Bak-Jensen, K.S., Mori, H., Juge, N., Nøhr, J., and Svensson, B. (2005). Involvement of individual subsites and secondary substrate binding sites in multiple attack on amylose by barley α-amylase. Biochemistry *44*, 1824–1832.

Kuriki, T. (2000). Catalytic mechanism of glycoenzymes. In: Glycoenzymes (Eds. M. Ohnishi, T. Hayashi, S. Ishijima, T. Kuriki), pp. 3–17, Japan Scientific Societies Press, Tokyo.

Kuriki, T., and Imanaka, T. (1999). The concept of the α-amylase family: structural similarity and common catalytic mechanism. J. Biosci. Bioeng. *87*, 557–565.

Laderman, K.A., Asada, K., Uemori, T., Mukai, H., Taguchi, Y., Kato, I., and Anfinsen, C.B. (1993). α-amylase from the hyperthermophilic archaebacterium *Pyrococcus furiosus* – cloning and sequencing of the gene and expression in *Escherichia coli*. J. Biol. Chem. *268*, 24402–24407.

Latorre-Garcia, L., Adam, A.C., Manzanares, P., and Polaina, J. (2005). Improving the amylolytic activity of *Saccharomyces cerevisiae* glucoamylase by the addition of a starch binding domain. J. Biotechnol. *118*, 167–176.

Lee, H.S., Kim, J.S., Shim, K.H., Kim, J.W., Park, C.S., and Park, K.H. (2005a). Quaternary structure and enzymatic properties of cyclomaltodextrinase from alkalophilic Bacillus sp. I-5. Biologia, Bratislava *60 (Suppl. 16)*, 73–77.

Lee, S.S., Yu, S., and Withers, S.G. (2003). Detailed dissection of a new mechanism for glycoside cleavage: α-1,4-glucan lyase. Biochemistry *42*, 13081–13090.

Lee, S.S., Yu, S., and Withers, S. G. (2005b). Mechanism of action of exo-acting α-1,4-glucan lyase: a glycoside hydrolase family 31 enzyme. Biologia Bratislava *60 (Suppl. 16)*, 137–148.

Levy, I., Paldi, T., and Shoseyov, O. (2004). Engineering a bifunctional starch-cellulose cross-bridge protein. Biomaterials *25*, 1841–1849.

Libessart, N., and Preiss, J. (1998). Arginine residue 384 at the catalytic center is important for branching enzyme II from maize endosperm. Arch. Biochem. Biophys. *360*, 135–141.

Lovering, A.L., Lee, S.S., Kim, Y.W., Withers, S.G., and Strynadka, N. C. J. (2005). Mechanistic and structural analysis of a family 31 α-glycosidase and its glycosyl-enzyme intermediate. J. Biol. Chem. *280*, 2105–2115.

MacGregor, E. A. (2005). An overview of clan GH-H and distantly-related families. Biologia Bratislava *60 (Suppl. 16)*, 5–12.

MacGregor, E.A., Janeček, S., and Svensson, B. (2001). Relationship of sequence and structure to specificity in the α-amylase family of enzymes. Biochim. Biophys. Acta *1546*, 1–20.

MacGregor, E.A., Jespersen, H.M., and Svensson, B. (1996). A circularly permuted α-amylase-type α/β-barrel structure in glucan-synthesizing glucosyltransferases. FEBS Lett. *378*, 263–266.

Machius, M., Wiegand, G., and Huber, R. (1995). Crystal structure of calcium-depleted *Bacillus licheniformis* α-amylase at 2.2 Å resolution. J. Mol. Biol. *246*, 545–559.

Machovič, M., and Janeček, S. (2003). The invariant residues in the α-amylase family: just the catalytic triad. Biologia, Bratislava *58*, 1127–1132.

Matsuura, Y. (2002). A possible mechanism of catalysis involving three essential residues in the enzymes of α-amylase family. Biologia Bratislava, *57 (Suppl. 11)*, 21–27.

Matsuura, Y., Kusunoki, M., Harada, W., and Kakudo, M. (1984). Structure and possible catalytic residues of Taka-amylase A. J. Biochem. Tokyo *95*, 697–702.

Mikami, B. (2000). Structure of β-amylase: X-ray crystallographic analysis. In: Glycoenzymes (Eds. M. Ohnishi, T. Hayashi, S. Ishijima, T. Kuriki), 55–81, Japan Scientific Societies Press, Tokyo

Mikami, B., Adachi, M., Kage, T., Sarikaya, E., Nanmori, T., Shinke, R., and Utsumi, S. (1999a). Structure of raw starch-digesting *Bacillus cereus* β-amylase complexed with maltose. Biochemistry *38*, 7050–7061.

Mikami, B., Degano, M., Hehre, E.J., and Sacchettini, J.C. (1994). Crystal structures of soybean β-amylase reacted with β-maltose and maltal: active site components and their apparent roles in catalysis. Biochemistry *33*, 7779–7787.

Mikami, B., Hehre, E.J., Sato, M., Katsube, Y., Hirose, M., Morita, Y., and Sacchettini, J. C. (1993). The 2.0 Å resolution structure of soybean β-amylase complexed with α-cyclodextrin. Biochemistry *32*, 6836–6845.

Mikami, B., Yoon, H.J., and Yoshigi, N. (1999b). The crystal structure of the sevenfold mutant of barley β-amylase with increased thermostability at 2.5 Å resolution. J. Mol. Biol. *285*, 1235–1243.

Mouille, G., Maddelein, M.L., Libessart, N., Talaga, P., Decq, A., Delrue, B., and Ball, S. (1996). Preamylopectin processing: a mandatory step for starch biosynthesis in plants. Plant Cell *8*, 1353–1366.

Murakami, T., Kanai, T., Takata, H., Kuriki, T., and Imanaka, T. (2006). A novel branching enzyme of the GH-57 family in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. J. Bacteriol. *188*, 5915–5924.

Nakai, H., Okuyama, M., Kim, Y.M., Saburi, W., Wongchawalit, J., Mori, H., Chiba, S., and Kimura, A. (2005). Molecular analysis of α-glucosidase belonging to GH-family 31. Biologia Bratislava *60 (Suppl. 16)*, 131–135.

Nakajima, R., Imanaka, T., and Aiba, S. (1986). Comparison of amino acid sequences of eleven different α-amylases. Appl. Microbiol. Biotechnol. *23*, 355–360.

Oh, B.H. (2003). The same group of enzymes with different names: cyclomaltodextrinases, neopullulanases, and maltogenic amylases. Biologia, Bratislava *58*, 299–305.

Ohdan, K., Kuriki, T., Takata, H., Kaneko, H., and Okada, S. (2000). Introduction of raw starch-binding domains into *Bacillus subtilis* α-amylase by fusion with the starch-binding domain of *Bacillus* cyclomaltodextrin glucanotransferase. Appl. Environ. Microbiol. *66*, 3058–3064.

Oslancova, A., and Janeček, S. (2002). Oligo-1,6-glucosidase and neopullulanase enzyme subfamilies from the α-amylase family defined by the fifth conserved sequence region. Cell. Mol. Life Sci. *59*, 1945–1959.

Oyama, T., Kusunoki, M., Kishimoto, Y., Takasaki, Y., and Nitta, Y. (1999). Crystal structure of β-amylase from *Bacillus cereus* var. mycoides at 2.2 Å resolution. J. Biochem. (Tokyo) *125*, 1120–1130.

Przylas, I., Tomoo, K., Terada, Y., Takaha, T., Fujii, K., Saenger, W., and Sträter, N. (2000). Crystal structure of amylomaltase from *Thermus aquaticus*, a glycosyltransferase catalysing the production of large cyclic glucans. J. Mol. Biol. *296*, 873–886.

Pujadas, G., and Palau, J. (1997). Anatomy of a conformational transition of β-strand 6 in soybean β-amylase caused by substrate (or inhibitor) binding to the catalytical site. Protein Sci. *6*, 2409–2417.

Pujadas, G., and Palau, J. (1999). TIM barrel fold: structural, functional and evolutionary characteristics in natural and designed molecules. Biologia, Bratislava *54*, 231–254.

Pujadas, G., and Palau, J. (2001). Evolution of α-amylases: architectural features and key residues in the stabilization of the $(\beta/\alpha)_8$ scaffold. Mol. Biol. Evol. *18*, 38–54.

Pujadas, G., Ramirez, F. M., Valero, R., Palau, J. (1996). Evolution of β-amylase: patterns of variation and conservation in subfamily sequences in relation to parsimony mechanisms. Proteins *25*, 456–472.

Reardon, D., and Farber, G.K. (1995). The structure and evolution of α/β barrel proteins. FASEB J. *9*, 497–503.

Rigden, D.J. (2002). Iterative database searches demonstrate that glycoside hydrolase families 27, 31, 36 and 66 share a common evolutionary origin with family 13. FEBS Lett. *53*, 17–22.

Robyt, J. F. (2005). Inhibition, activation, and stabilization of α-amylase family enzymes. Biologia, Bratislava *60 (Suppl. 16)*, 17–26.

Rodriguez-Sanoja, R., Oviedo, N., and Sanchez, S. (2005). Microbial starch-binding domain. Curr. Opin. Microbiol. *8*, 260–267.

Sinnot, M.L. (1990). Catalytic mechanisms of enzymatic glycosyl transfer. Chem. Rev. *90*, 1171–1202.

Søgaard, M., Kadziola, A., Haser, R., and Svensson, B. (1993). Site-directed mutagenesis of histidine 93, aspartic acid 180, glutamic acid 205, histidine 290, and aspartic acid 291 at the active site and tryptophan 279 at the raw starch binding site in barley α-amylase 1. J. Biol. Chem. *268*, 22480–22484.

Stoffer, B., Aleshin, A.E., Firsov, L.M., Svensson, B., and Honzatko, R. B. (1995). Refined structure for the complex of D-gluco-dihydroacarbose with glucoamylase from *Aspergillus awamori* var. X100 to 2.2 Å resolution: dual conformations for extended inhibitors bound to the active site of glucoamylase. FEBS Lett. *358*, 57–61.

Svensson, B. (1994). Protein engineering in the α-amylase family: catalytic mechanism, substrate specificity, and stability. Plant Mol. Biol. *25*, 141–157.

Svensson, B., Jensen, M.T., Mori, H., Bak-Jensen, K.S., Bønsager, B., Nielsen, P.K., Kramhøft, B., Prætorius-Ibba, M., Nøhr, J., Juge, N., Greffe, L., Williamson, G., and Driguez, H. (2002). Facinating facets of function and structure of amylolytic enzymes of glycoside hydrolase family 13. Biologia, Bratislava *57 (Suppl. 11)*, 5–19.

Svensson, B., Jespersen, H., Sierks, M.R., and MacGregor, E.A. (1989). Sequence homology between putative raw starch-binding domains from different starch degrading enzymes. Biochem. J. *264*, 309–311.

Sevcik, J., Solovicova, A., Hostinova, E., Gasperik, J., Wilson, K.S., and Dauter, Z. (1998). Structure of glucoamylase from *Saccharomycopsis fibuligera* at 1.7 Å resolution. Acta Crystallogr. D Biol. Crystallogr. *54*, 854–866.

Totsuka, A., and Fukazawa, C. (1996). Functional analysis of Glu380 and Leu383 of soybean β-amylase. A proposed action mechanism. Eur. J. Biochem. *240*, 655–659.

Tranier, S., Deville, K., Robert, X., Bozonnet, S., Haser, R., Svensson, B., and Aghajari, N. (2005). Insights into the "pair of sugar tongs" surface binding site in barley α-amylase isozymes and crystallization of appropriate sugar tongs mutants. Biologia, Bratislava *60 (Suppl. 16)*, 37–46.

Turner, P., Nilsson, C., Svensson, D., Holst, O., Gorton, L., and Nordberg Karlsson, E. (2005). Monomeric and dimeric cyclomaltodextrinases reveal different modes of substrate degradation. Biologia, Bratislava *60 (Suppl. 16)*, 79–87.

Van Lieshout, J.F.T., Verhees, C.H., Ettema, T.J.G., van der Saar, S., Imamura, H., Matsuzawa, H., van der Oost, J., and de Vos, W.M. (2003). Identification and molecular characterization of a novel type of α-galactosidase from *Pyrococcus furiosus*. Biocatal. Biotransform. *21*, 243–252.

Vihinen, M., and Mäntsala, P. (1989). Microbial amylolytic enzymes. Crit. Rev. Biochem. Mol. Biol. *24*, 329–418.

Vihinen, M., Ollikka, P., Niskanen, J., Meyer, P., Suominen, I., Karp, M., Holm, L., Knowles, J., and Mäntsälä, P. (1990). Site-directed mutagenesis of a thermostable α-amylase from *Bacillus stearothermophilus*: putative role of three conserved residues. J. Biochem. *107*, 267–272.

Yu, S., Bojsen, K., Svensson, B., and Marcussen, J. (1999). α-1,4-glucan lyases producing 1,5-anhydro-D-fructose from starch and glycogen have sequence similarity to α-glucosidases. Biochim. Biophys. Acta *1433*, 1–15.

Zona, R., Chang-Pi-Hin, F., O'Donohue, M.J., and Janeček, S. (2004). Bioinformatics of the glycoside hydrolase family 57 and identification of catalytic residues in amylopullulanase from *Thermococcus hydrothermalis*. Eur. J. Biochem. *271*, 2863–2872.

CHAPTER 2

THE USE OF STARCH PROCESSING ENZYMES IN THE FOOD INDUSTRY

JÓZEF SYNOWIECKI*

Department of Food Chemistry, Technology and Biotechnology, Chemical Faculty, Gdansk University of Technology, Gdansk, Poland

**synowiec@chem.pg.gda.pl*

1. INTRODUCTION

Starch, the main component of many agricultural products, *e.g.* corn (maize), potatoes, rice and wheat, is deposited in plant cells as reserve material for the organism in the form of granules which are insoluble in cold water. This carbohydrate is the main constituent of food products such as bread and other bakery goods or is added to many foods for its functionality as a thickener, water binder, emulsion stabilizer, gelling agent and fat substitute. Starch granules consist of two types of molecules composed of α-D-glucose units called amylose and amylopectin. In amylose almost all the glucose residues are linked by α-1,4-glycosidic bonds, whereas in amylopectin about 5 % of the carbohydrate units are also joined by α-1,6-linkages forming branch points. The relative contents of amylose and amylopectin depend on the plant species. For example, wheat starch contains about 25% amylose while waxy corn starch is more than 97–99% amylopectin. Starch origin also makes differences to the size, shape and structure of the polysaccharide granules, their swelling power, gelatinisation temperature, extent of esterification with phosphoric acid, and the amounts of lipids and other compounds which are retained inside the hydrophobic inner surface of the amylose helices.

Expanding starch functionality can be achieved through chemical or enzymatic modifications. The most important methods of enzymatic starch processing (Fig. 1) are the production of cyclodextrins and the hydrolysis of starch into a mixture of simpler carbohydrates for the production of syrups having different compositions and properties. These products are used in a wide variety of foodstuffs: soft drinks, confectionery, meats, packed products, ice cream, sauces, baby food, canned fruit,

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 19-34

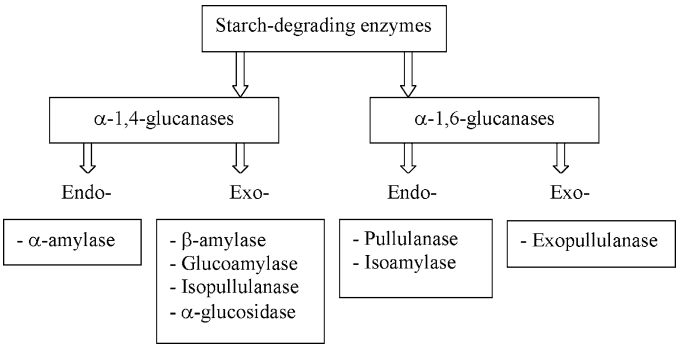


Figure 1. Starch degrading enzymes

preserves, *etc*. Furthermore, glucose produced during starch hydrolysis can be converted to fuel alcohol and other bio–products by yeast or bacterial fermentation, or isomerised to fructose in a reaction catalysed by glucose isomerase. High fructose syrup is used as a sweetener in different food products and is more suitable for diabetics than ordinary household sugar.

2. ENZYMES USED FOR STARCH HYDROLYSIS

2.1. α-Amylases

The industrial degradation of starch is usually initiated by α-amylases (α-1,4-glucanohydrolases) a very common enzyme in micro-organisms. Together with other starch-degrading enzymes (eg. pullulanases), α-amylases are included in family 13 of glycosyl hydrolases (Henrissat and Bairoch, 1996) characterized by a $(\alpha/\beta)_8$-barrel conformation (Fig. 2A). The structural and functional aspects of α-amylases have been reviewed by Nielsen and Borchert (2000) and MacGregor *et al.* (2001). The enzyme contains a characteristic substrate binding cleft (Fig. 2B) that can accommodate between four to ten glucose units of the substrate molecule. Each binding site has affinity to only one glucose unit of the carbohydrate chain. However, the interactions of oligosaccharides with several binding sites creates a multipoint linkage which results in the correct arrangement of long substrate molecules towards the catalytic site. Differences in the number of substrate binding sites and the location of catalytic regions determine substrate specificity, the length of the oligosaccharide fragments released after hydrolysis and the carbohydrate profile of the final product. Substrate binding is not sufficient for catalysis when all the glucose residues of the engaged oligosaccharide chain fall outside the catalytic region (Fig. 2C). This phenomenon occurs only in cases of advanced hydrolysis producing oligosaccharide molecules which are too short to occupy all the substrate binding sites. The probability of inappropriate binding contributes to a rapid decrease in the reaction rate during the final stages of reaction and also

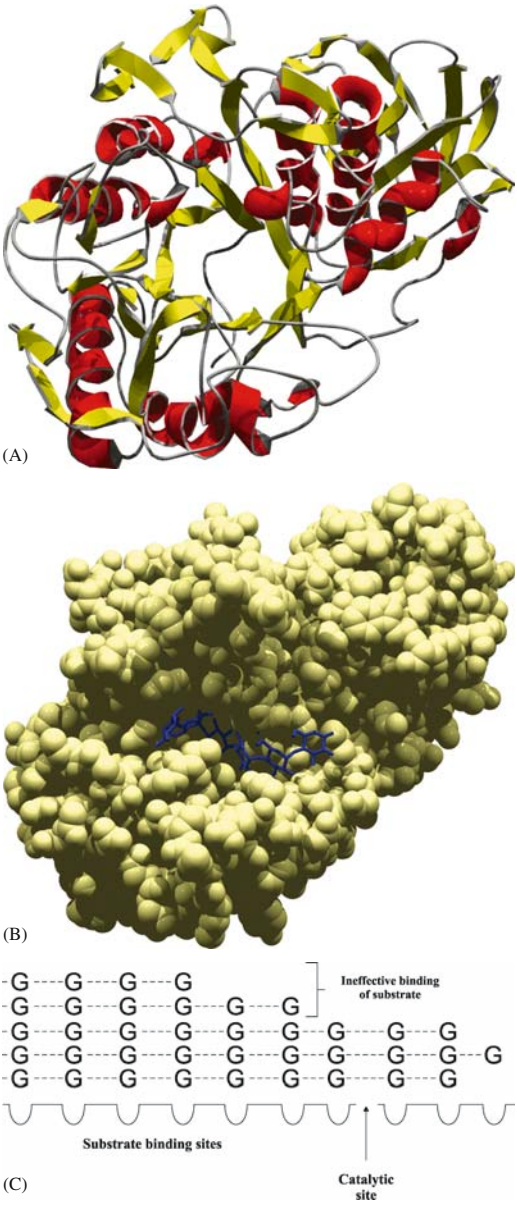


Figure 2. Structure of α-amylases. **A**: Overal structure of porcine pancreatic α-amylase, a representative member of family 13 glycosyl hydrolases. **B**: Visualization of the inhibitory oligosaccharide V-1532 bound to the catalytic cleft of the same enzyme (Machius *et al*, 1996). **C**: Schematic representation of the catalytic cleft. G represents the glucose units of the substrate

explains differences in the carbohydrate profiles of the final products generated by α-amylases originating from various sources. Other domains in the α-amylase molecule maintain the structure of the protein. One of these called "the starch-binding domain" has affinity for starch granules in those enzymes which can degrade starch without the necessity for its gelatinisation. All structural differences result in a great diversity in enzyme activity, stability, reaction conditions and substrate specificity, which vary both in preference for chain length and the ability to cleave the α-1,4-bonds close to the α-1,6-branch point in amylopectin molecules. For example, the temperature-activity optima of microbial α-amylases range from approximately 25 °C to 95 °C. Calcium ions play a significant role in maintaining the structural integrity of the catalytic and/or substrate binding sites in α-amylases, amylopullulanases and several other glycosyl hydrolases. Thus the addition of calcium salts to the reaction mixture essentially improves enzyme activity and stability. Nevertheless, excessive amounts of Ca^{2+} induce inhibitory effects and decrease the reaction yield.

α-Amylases catalyse cleavage of α-1,4-glycosidic bonds in the inner region of the molecule hence causing a rapid decrease in substrate molecular weight and viscosity. These endo-acting enzymes can be divided into liquefying or saccharyfying α-amylases which preferentially degrade substrates containing more than fifteen or four glucose units, respectively. Prolonged hydrolysis of amylose leads to carbohydrate conversion into maltose, maltotriose and oligosaccharides of varying chain lengths, sometimes followed by a second stage in the reaction releasing glucose from maltotriose. However, the reaction rate is diminished when the enzyme acts on small oligosaccharide molecules. Some α-amylases, *e.g.* that from *Pyrococcus furiosus*, cannot release glucose because maltopentaose is the smallest substrate hydrolysed by this enzyme (Dong *et al*, 1997). Hydrolysis of amylopectin or glycogen also yields glucose, maltose and maltooligosaccharides in addition to a series of branched "α-limit dextrins" containing four or more glucose residues in the neighbourhood of an α-1,6-glycosidic bond originating from branch points in the polysaccharide molecule. During the hydrolysis catalysed by these enzymes the hydroxyl groups formed during cleavage of the glycosidic bonds retain the α-configuration while β-amylase and glucoamylase, belonging to other enzyme families, cause inversion to the anomeric β-configuration (Janeček, 1997).

α-Amylases are used in a number of industrial processes which take place under diverse physical and chemical conditions. Thus, for each individual application the enzyme which best meets the particular demands of the process is desirable. High thermostability is sometimes desired because elevated temperatures improve starch gelatinisation, decrease media viscosity, accelerate catalytic reactions and decrease the risks of bacterial contamination. An additional benefit of high-temperature catalysis is the inactivation of enzymes originating from food materials which give rise to undesirable reactions during processing. The most thermostable α-amylase currently used in biotechnological processes is produced by*Bacillus licheniformis*. It remains active for several hours at temperatures over 90 °C under conditions of industrial starch hydrolysis. A potential source of α-amylases functioning at

even higher temperatures are hyperthermophilic archaea. The extracellular enzyme of *Pyrococcus woesei* is active between 40 °C and 130 °C with an optimum at 100 °C and pH 5.5 (Koch *et al*, 1991). The intracellular α-amylase from a related species, *Pyrococcus furiosus*, exhibits maximal activity at the same temperature but the optimum pH is 6.5–7.5 (Ladermann *et al*, 1993). To inactivate the enzyme from *Pyrococcus woesei* completely, autoclaving at 120 °C for 6 h is necessary. However, for industrial starch processing α-amylases retaining high activity at pH around 4.0 are desired. None of the most thermostable α-amylases have high stability at this pH, therefore protein engineering studies concerning improvement of this property have been initiated. By contrast, the thermolabile α-amylases are usually used for starch saccharification at moderate temperatures, *e.g.* in the brewing industry, the preparation of fermentation broth in alcohol distilleries, in dough conditioning or as a detergent additive.

2.2. Debranching Enzymes

There are two main groups of endo-acting debranching enzymes which can cleave the α-1,6-glycosidic linkages existing at the branch points of amylose, glycogen, pullulan and related oligosaccharides. The first group are pullulanases that specifically attack α-1,6- linkages, liberating linear oligosaccharides of glucose residues linked by α-1,4- bonds. The second group of debranching enzymes are neopullulanases and amylopullulanases, which are active toward both α-1,6- and α-1,4- linkages.

Pullulanases are generally produced by plants, *e.g.* rice, barley, oat and bean, as well as by mesophilic micro-organisms such as: *Klebsiella*, *Escherichia*, *Streptococcus*, *Bacillus* and *Streptomyces*. These enzymes are rather heat-sensitive, and commercially available preparations obtained from *Klebsiella pneumoniae* or *Bacillus acidopullulyticus* should be used at temperatures not exceeding 50–60 °C. Nevertheless, the search for efficient sources of thermostable debranching enzymes is underway because the enzymatic conversion of starch is usually carried out at elevated temperatures. Pullulanases are seldom produced by thermophiles. However, a recent study shows that a good source of heat-resistant pullulanase is the aerobic, thermophilic bacterium *Thermus caldophilus* which syntheses an enzyme that is optimally active at 75 °C and pH 5.5 and retains activity up to 90 °C (Kim *et al*, 1996).

Most of the heat-resistant debranching enzymes belong to the group of amylopullulanases which are widely distributed among thermophilic bacteria and archaea, and have been isolated from cultures of *Bacillus subtilis*, *Thermoanaerobium brockii*, *Clostridium thermosulphuricum* and *Thermus aquaticus* (Ara *et al*, 1995). The enzyme from *Pyrococcus woesei* which displays maximal activity at 105 °C and pH 6.0 is the most thermostable amylopullulanase known and has been purified and expressed in *Eschericha coli* (Leuschner and Antranikian, 1995). Thermostable amylopullulanases should be valuable components of laundry and dishwashing detergents since they catalyse both debranching as well as liquefying reactions. However, their applications are limited because amylopullulanases of bacterial origin are seldom active at alkaline pH.

2.3. Exo-acting Amylases

Two types of exo-acting hydrolases are commonly used for starch saccharification: β-amylases (EC 3.2.1.2) and glucoamylases (EC 3.2.1.3). Both act on glycosidic linkages at the non-reducing ends of amylose, amylopectin and glycogen molecules, producing low-molecular weight carbohydrates in the β-anomeric form. The main end-product of hydrolysis catalysed by β-amylases is maltose, while glucoamylase (amyloglucosidase) generates glucose. Structurally, β-amylases and glucoamylases are included in families 14 and 15 of the classification of Henrissat and Bairoch (1996), respectively. Whereas β-amylases present an $(\alpha/\beta)_8$ fold similar to α-amylases, glucoamylases are characterized by an $(\alpha/\alpha)_6$ structure.

All β-amylases are unable to cleave α-1,6-linkages and the final product consists of maltose and "β-limit dextrin". Thus degradation of amylopectin is incomplete, resulting in only 50–60 % conversion to maltose. Even in the case of amylose, the maximum degree of hydrolysis is 75–90 % because this polysaccharide also has a slightly branched structure. Accumulation of "β-limit dextrin" is undesirable because it increases the viscosity of maltose syrups. β-Amylases occur in higher plants, such as barley, wheat, sweet potatoes and soybeans and have also been discovered in strains of *Pseudomonas*, *Bacillus*, *Streptococcus* and some other micro-organisms. These enzymes are rare among thermophiles, and currently produced β-amylases are not stable at temperatures above 60 °C. Application of more heat-resistant enzymes which are active in slightly acidic environments will reduce saccharification time and can limit the risk of unwanted browning reactions at alkaline or neutral pH values. Shen and co-workers (1988) reported that β-amylase from *Clostridium thermosulfurigenes* is an option, since it displayed maximal activity at 75 °C and exhibits broad pH stability over the range 4.0 to 7.0.

Glucoamylases cleave preferentially α-1,4-linkages and can also cleave α-1,6-glycosidic linkages, although at a much lower rate. As a consequence, glucoamylases have the ability to carry out almost complete degradation of starch into glucose. At concentrations of glucose in reaction media exceeding 30–35 % the glucoamylases can catalyse the reverse reactions forming maltose, isomaltose and other by-products thereby decreasing the final yield of the process. Glucoamylases are widely distributed among plants, animals and mesophilic micro-organisms, such as *Saccharomyces*, *Endomycopsis*, *Aspergillus*, *Penicillium*, *Mucor* and *Clostridium*. Generally, the enzymes from these sources exhibit the highest activity at temperatures ranging from 45°C to 60°C and at pH 4.5 to 5.0. Like β-amylases, glucoamylases are rare among thermophiles.

3. ENZYMATIC PROCESSING OF STARCH AND STARCH-CONTAINING FOOD

3.1. Products Obtained During Starch Hydrolysis

Starch hydrolases are important industrial enzymes which are used as additives in detergents, for the removal of starch sizing from textiles, the liquefaction of

starch and the proper formation of dextrins in baking. They are also added to break down the starch that accompanies saccharose in sugar cane juice and interferes with filtration. The discovery and application of enzymes exhibiting different activities and substrate specificities isolated from a variety of microbial sources or obtained by gene cloning or protein engineering has resulted in the development of many starch products of diverse carbohydrate profiles and functional properties. The hydrolysis products obtained are usually divided in two main groups characterized by low- or high-degrees of starch conversion. In the first group are those maltodextrins prepared by limited hydrolysis (DE 10–20) of gelatinised starch in reactions commonly catalysed by heat-resistant α-amylases, without subsequent saccharification. Maltodextrins provided for some applications are additionally processed by debranching enzymes to remove the side chains of amylopectin molecules thus producing linear oligosaccharides. The main components of these products, found in amounts ranging 75–96 % of dry weight, are oligosaccharides containing more than four glucose residues. Maltodextrins have useful functional properties, *e.g.* low hygroscopicity, high solution viscosity, low sweetness as well as the ability to retard ice crystal growth in ice-cream and other frozen foods. These attributes make them suitable for the formulation of different coatings, improvement of the chewiness and binding properties of food products, and for moisture retention in soft or hard candies. Maltodextrins also have applications as binders for encapsulated pharmaceuticals, the protection of encapsulated flavours from oxidation, or as lipid substitutes in low-fat food products. For these purposes starch syrups with higher degrees of hydrolysis (DE 20–70) and containing 40–78 % of oligosaccharides larger than maltotetraose can also be used. These hydrolysates are available in the form of viscous solutions and increase the resistance of starch gels to retrogradation and prevent the crystallization of sucrose. They are often exploited as thickeners in many food products.

Advanced starch hydrolysis which leads to products including significant amounts of maltose and glucose can be achieved during prolonged (48–96 h) times of saccharification. Maltose is the main component of the hydrolysates called high-maltose-, extremely high-maltose- and high-conversion syrups, containing on a dry basis 35–40 %, 70–85 % and 30–47 % of this carbohydrate, respectively. High-conversion syrups also contain large amounts of glucose, ranging from 35 % up to 45 % on a dry basis. Hydrolysates containing maltose are usually exploited as sweeteners, flavour and taste enhancers, moisture conditioners, stabilizers to protect against the crystallization of sucrose in confectioneries as well as a cryoprotectant controlling ice crystal formation in frozen food. The high-maltose syrups have low viscosity and hygroscopy, mild sweetness and reduced browning capacity during heating. These products are also used to replace sucrose in foods for diabetics and for the synthesis of maltulose or maltitol which are utilized as low-calorie sweeteners. Other recently developed applications for maltose syrups or maltooligosaccharide solutions obtained during starch processing are the production of trehalose and cyclodextrins.

A characteristic property of high-glucose syrups is their participation and intensification of Maillard reactions, developing the desired flavours and brown colour of

fried or baked goods. Besides applications as food additives, glucose syrups are also converted into fructose. The isomerisation efficiency depends on the glucose content of the substrate. Theoretically, glucoamylase can completely hydrolyse amylose to glucose but a limited level of glucose in the final product is caused by maltulose (4-α-D-glucopyranosyl-D-fructose) synthesis and by reverse reactions which lead to the formation of maltose, isomaltose and α-1,6-oligosaccharides. Maltulose is accumulated in the product because glucoamylases do not cleave the glycosidic bonds between glucose and fructose residues. Undesirable maltulose synthesis can be eliminated when saccharification is catalysed at pH below 6.0.

3.2. Production of starch hydrolysates

There are two basic steps in the enzymatic conversion of starch (see Fig. 3): liquefaction and saccharification. During liquefaction the concentrated slurry of starch granules (30–40 %, w/v) is gelatinised at an elevated temperature (90–110°C). The addition of thermostable endoamylase (EC 3.2.1.1) at this stage of the process protects against a rapid increase in starch solution viscosity caused by the release of amylose from swelling starch granules (Guzman-Maldonato and Paredes-Lopez, 1995). Enzymatic hydrolysis of amylose by α-amylase proceed until the chain lengths of the reaction products are about 10–20 glucose units. At this point the starch fragments fail to bind well to the enzyme. Hydrolysis of amylopectin results not only in the production of a mixture of linear maltooligosaccharides, as does amylose hydrolysis, but also fragments that contain the α-1,6- bond which cannot be cleaved by α-amylase. Studies have been done on the immobilization of α-amylase on different supports (Synowiecki *et al*, 1982; Lai *et al*, 1998). However, the reaction rate was found to be strongly influenced by diffusion limitations caused by the high molecular weight of the substrate and high solution viscosity. Other glucosyl hydrolases that do not act on starch but yield improvements in starch processing are xylanases and cellulases. Both are involved in the cleavage of the β-1,4-glycosidic bonds linking residues of D-glucose or D-xylopyranose in cellulose and xylans, respectively. Xylanases reduce the viscosity of wheat starch slurry by degrading arabinoxylans and other xylans, whereas cellulases positively affect the filterability of the final products of starch hydrolysis in the case of its contamination by cellulose fibres.

The saccharification step is carried out at a lower temperature and leads to the hydrolysis of the oligosaccharides obtained into glucose or maltose in reactions catalysed by glucoamylase (EC 3.2.1.3) or β-amylase (EC 3.2.1.2), respectively. The yield of starch hydrolysis may be enhanced by using glucoamylase or β-amylase in combination with pullulanase (EC 3.2.1.41) or other debranching enzymes. In general the use of pullulanase increases the glucose yield up to 94 % (Crabb and Mitchinson, 1997).

Since the gelatinisation of starch granules is completed near 100°C in the majority of industrial processes, thermostable α-amylases are used. These enzymes are widespread among thermophilic bacteria and archea, and the genes encoding a few

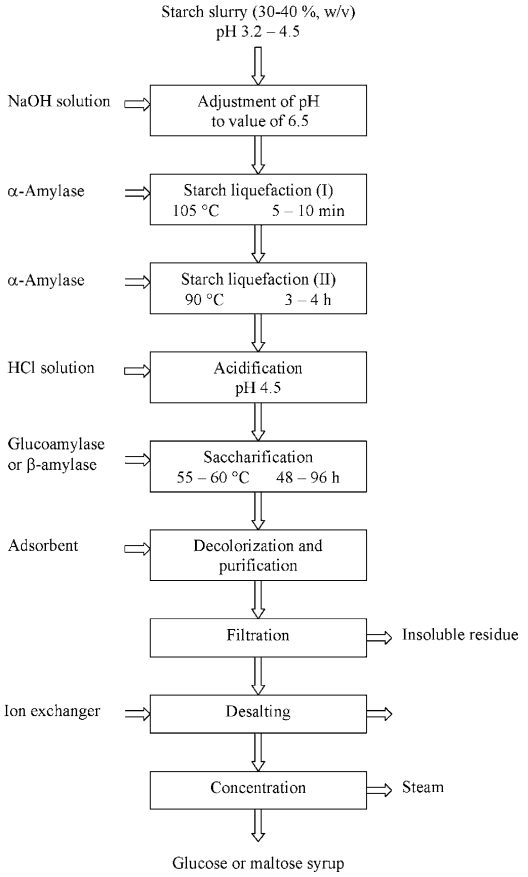


Figure 3. Flowsheet for glucose or maltose syrup production

of them have been cloned and expressed in mesophilic hosts (Frillingos *et al*., 2000; Grzybowska *et al*., 2004). Termamyl originates from *Bacillus licheniformis*, and other α-amylase preparations used for starch liquefaction usually show highest activity at temperatures above 90 °C and at pH 5.5 to 6.0. These conditions are not however compatible with those of the glucoamylases or β-amylases used in the next step which are more sensitive to heat and are inactivated above 60 °C. Limited enzyme thermostability implies that rapid cooling of the substrate is required before further processing can proceed but this leads to an increase in the viscosity of the reaction mixture and a decrease in the final yield of the process. Since the natural pH of starch slurry is approximately 4.5 it should be adjusted to the value desirable for maximal enzyme activity during substrate liquefaction and then reduced to 4.5 prior to the saccharification step. The necessity for temperature and pH adjustments

increases the costs of the process and requires additional ion-exchange refinement of the final product for removal of the NaCl synthesised.

An important development would be to carry out starch degradation in a single step. This can be achieved using more heat-resistant α-amylases which can operate at lower pH values than the enzyme from *Bacillus licheniformis* and do not require calcium salts for activity. Improved thermostability avoids the need for further addition of α-amylase during liquefaction to replace that destroyed by high-temperature treatment. α-Amylases from different thermophiles show promising properties, but none has yet been produced on a commercial scale. For further oligosaccharide depolymerisation enzymes catalysing saccharification under conditions compatible with those used for α-amylase activity are necessary. This would make possible the application of all the enzymes together without the need for temperature and pH adjustments before liquefaction and saccharification. Recent investigations show that the oligosaccharides released during prolonged α-amylase action on starch can be hydrolysed by thermostable α-glucosidases (EC 3.2.1.20). These enzymes act on terminal non-reducing α-1,4- and to a lesser extent, α-1,6-glucosidic linkages, forming glucose as an end-product. Most of the α-glucosidases obtained from thermophiles and mesophiles showed greatest activity towards maltose and isomaltose (Kelly and Fogarty, 1983). However, significant activity against maltooligosaccharides makes these enzymes suitable for use in the last step of starch degradation instead of the more heat sensitive glucoamylases. Especially suitable are those α-glucosidases with increased ability to hydrolyse the α-1,6-glucosidic bonds occurring at the branch points of the amylopectin molecule. Legin and co-workers (1998) demonstrated the feasibility of glucose syrup production using thermostable α-glucosidase from *Thermococcus hydrothermalis* in cooperation with α-amylases and pullulanases. We have reported that an alternative source of thermostable enzyme having α-glucosidase activity is the halotolerant, non-sporulating bacterium *Thermus thermophilus* from marine and terrestrial hot springs (Zdzieblo and Synowiecki, 2002). The half-life of this enzyme incubated at 85 °C is about 2h, and at 95 °C no measurable activity remains after 30 min. The application of "thermozymes" for starch saccharification increases the conversion yield, enhances solubility and decreases the viscosity of the substrate solution. Moreover, the low levels of activity of thermostable enzymes at reduced temperatures facilitate the termination of the reaction simply by cooling. An alternative to starch processing using thermostable α-amylase is the application of endo-glucanase which has activity towards native starch granules, as for example glucoamylase from *Rhizopus* sp. (James and Lee, 1997).

3.3. Glucose Isomerisation

The isomerisation of starch-derived glucose to fructose leads to greater sweetness of the obtained syrup which is commonly used in many food and beverage products, *e.g.* as a sweetener and an enhancer of citrus flavour. Fructose is the sweetest tasting of all the carbohydrates and is suitable for the formulation of low-calorie products

having reduced sucrose content, or as a sweetener for diabetics because it can be metabolised without insulin. The use of fructose syrup as an additive to some baked products results in desirable browning developed as a result of Maillard reactions. In addition, fructose acts as a crystallization inhibitor which keeps sucrose in solution thus producing a cookie that retains its soft texture during storage.

Fructose syrups are usually made in a continuous process catalysed by immobilized glucose (xylose) isomerase (EC 5.3.1.5) at temperatures of 55–60 °C. Under these conditions only 40-42 % of the glucose is converted to fructose. The process yield can be enhanced at higher temperatures which shifts the equilibrium of the isomerisation towards increased fructose concentrations. However, this limits the half-life of the enzyme obtained from mesophilic sources and increases the amount of by-products created by the Maillard reactions that occur at the slightly alkaline pH values necessary for maximum activity of glucose isomerase. In order to produce the syrup containing the standard concentration (55 %) of fructose, cation-exchange fractionation of carbohydrates is used (Crabb and Mitchinson, 1997). During this step fructose is retained on the chromatographic matrix while glucose and higher saccharides pass through the column and are returned to the isomerisation unit. The adsorbed fructose is then released by elution with water and the eluate contains more than 90 % fructose on a dry basis. The product is then mixed with 42 % fructose syrup to the final concentration required for many applications. This chromatographic step can be omitted when glucose conversion is catalysed by more efficient thermostable glucose isomerase having increased activity at the acidic pH values necessary for reducing undesirable side reactions. Since glucose isomerases active at elevated temperatures are synthesised by various species of *Thermus* and some other thermophilic micro-organisms, future industrial application of these enzymes will lead to significant reductions in production costs (Vieille and Zeikus, 2001).

3.4. Trehalose Production

Starch or maltose syrups can be successfully processed into trehalose in reactions catalysed by enzymes isolated from mesophilic or thermophilic micro-organisms. Trehalose (α-D-glucopyranosyl α-D-glucopyranoside) is a stable, non-reducing disaccharide containing 1,1 glycosidic linkages between the glucose moieties. This carbohydrate is involved in protection of biological structures during freezing, desiccation or heating (Richards *et al.*, 2002). Amorphous glass trehalose holds trapped biological molecules without introducing changes in their native structure and consequently limits the damage inflicted on biological materials during desiccation. Furthermore, this non-hygroscopic glass is permeable to water but impermeable to hydrophobic, aromatic esters. It minimizes the undesirable loss of hydrophobic flavour compounds and thus facilitates the production of dried foods retaining the aroma similar to the fresh product. Trehalose can be used in the food, cosmetics, medical and biotechnological industries, and as stabilizer of vaccines, enzymes, antibodies, pharmaceutical preparations and organs for transplantation. The mild sweetness of trehalose, its low cariogenicity, good solubility in water, stability under

low pH conditions, reduction of water activity, low hygroscopicity, depression of freezing point, high glass transition temperature and protein protection properties make it a valuable food ingredient. This compound does not caramelise and does not undergo Maillard reactions, it is safe for human consumption and has been accepted by the European regulatory system (Richards *et al*, 2002). Trehalose may be used in wide range of products including beverages, chocolate and sugar confectionery, bakery, dairy and fruit products and as a cryoprotectant for surimi and other frozen foods.

Trehalose can be produced from starch by using two novel enzymes derived from certain mesophiles, *e.g. Arthrobacter*, *Brevibacterium*, *Micrococcus* and *Rhizobium*, as well as from the hyperthermophilic archaeon *Sulfolobus shibatae* (Lama *et al*, 1990; Nakada *et al*, 1996; Di Lernia *et al*, 1998). These enzymes are designated as maltooligosyl-trehalose synthase and maltooligosyl-trehalose trehalohydrolase. The former converts the terminal α-1,4-linkage at the reducing end of the maltooligosaccharide molecule to the α, α-1,1-bond existing in trehalose; the latter releases trehalose during hydrolysis of α-1,4-linkage between the second and third glucose units, and this reaction repeats until the remaining oligosaccharide consists of no more than two or three glucose units. The maltooligosaccharides used as substrates for these reactions are produced by treatment of liquified starch slurry by debranching enzymes.

Other sources of enzymes for trehalose synthesis are micro-organisms containing trehalose synthase (EC 5.4.99.16). This enzyme catalyses intramolecular transglucosylation and leads to conversion of the α-1,4- glucosidic linkage of maltose into α, α-1,1-bonds (Nishimoto *et al*, 1996). As a result, maltose is converted into trehalose, producing a small amount of glucose as a by-product. A conversion yield reaching 80 % or more indicates the suitability of trehalose synthase for the industrial production of trehalose from maltose syrup in a one-step process. Trehalose synthase is produced by *Pimelobacter* sp. and a few other mesophiles. However, the thermostable enzyme, *e.g.* from *Thermus caldophilus* with optimum activity at 65 °C, seems to be more suitable because the higher conversion temperature prevents contamination of the reaction mixture by micro-organisms.

3.5. Cyclodextrin Synthesis

Starch degrading enzymes are also used for the production of cyclodextrins. In the first stage of this process both α-amylases and pullulanases are involved in creating unbranched oligosaccharides. Subsequently, the resulting linear molecules are cleaved by cyclomaltodextrin glucanotransferase, and enzyme first isolated from *Bacillus macerans*, to yield oligosaccharides of 6-8 units. As a consequence of the helical structure of these oligosaccharides, the two ends of each molecule are in close proximity to each other, therefore they are easily joined together to form the ring structure characteristic of cyclodextrins. The final product is a mixture of α-, β- and γ-cyclodextrins, composed of six, seven or eight α-1,4-linked glucose residues. The proportion of each type can be controlled through enzyme selectivity as well

as by the temperature and pH of the reaction media. In some production methods, the selectivity of the process is improved when the substrate solution contains an appropriate organic solvent which directs the reaction to produce only one type of cyclodextrin (Guzman-Maldonado and Paredes-Lopez, 1995). The product precipitates in the form of an insoluble complex, decanol or cyclooctane being used for the preparation of α- or β-cyclodextrins respectively, and it is separated by centrifugation or filtration. Cyclodextrin production without the application of solvents may lead to microbial contamination but this problem can be prevented by raising the reaction temperature. Recently, a heat-resistant cyclomaltodextrin glucanotransferase was found in *Thermococcus* species (Viele and Zeikus, 2001). This enzyme is very stable at temperatures up to 100 °C and also possesses α-amylase activity. This property allows the production of cyclodextrins without the need for addition of α-amylase for preliminary starch liquefaction.

The hydroxyl groups of a cyclodextrin molecule are located on the surface of the oligosaccharide ring, whereas its interior is apolar and can easily form inclusion complexes with hydrophobic compounds of adequate size and structure. This property makes cyclodextrins suitable for many applications in the food, cosmetics and pharmaceutical industries, since they can capture undesirable tastes or odours, stabilize volatile compounds and increase the solubility of hydrophobic substances in water. For example, cyclodextrins are used for the debittering of citrus juices, protecting lipids against oxidation or for the removal of cholesterol from eggs (Shaw *et al*, 1984; Szejtli, 1982).

3.6. Significance of Amylolytic Enzymes in Food Processing

Starch hydrolysing enzymes play a significant role in the processing of some raw food materials, especially in the baking and brewing industries as well as in the production of soft and alcoholic drinks. The enzymes necessary for these purposes are often natural components of raw food materials, *e.g.* α- and β-amylases in flour, or are sourced from malt or other preparations obtained from higher plants and micro-organisms.

In the baking industry, the α- and β-amylases of the cereal grain play an essential role. However, their content in flour depends on the climatic conditions during ripening and harvesting. When the weather is very humid the grain starts to germinate and the content of amylolytic enzymes is too high for the preparation of good quality bakery goods. In contrast, the flour obtained from cereals cultivated in a hot and dry climate often has a very low α-amylase content and its deficit needs to be supplemented. The α- and β-amylases have different but complementary functions during the bread making process (Martin and Hoseney, 1991). The α-amylases break down starch into low-molecular weight dextrins. β-amylase converts these oligosaccharides into maltose which is necessary for yeast growth. Insufficient amounts of fermentable sugars diminish the secretion of carbon dioxide leading to limited dough rise and decreased crumb volume.

Appropriate levels of amylolytic enzymes are especially important during bread making for the formation of dextrins which contribute to the browning of the crust, add flavour to the bread as well as influencing the degree of staling retardation. Staling is mainly caused by starch retrogradation leading to limited water holding capacity and reduced crumb elasticity. The mechanism of starch retrogradation is still not well understood. However, it is known that susceptibility to retrogradation depends on the amount of linear amylose present in starch and can be diminished when the side chains of branched amylopectin molecules are shortened by the action of maltogenic α-amylases (Christophersen and Otzen, 1998). Excessive levels of dextrin formation, often causing collapse of the bread after baking, leads to a final product with an unacceptable gummy and sticky structure. Thus dextrin should be formed only at the beginning of baking when the dough is placed in the oven, the enzyme efficiency steadily increasing with the rise in temperature until its thermal inactivation. Most α-amylases, *e.g.* that from barley malt and the commonly used enzyme from *Aspergillus niger*, have limited anti-staling effects due to their inactivation prior to starch gelatinization. This inconvenience can be avoided by the use of heat-resistant substitutes of mesophilic α-amylases, *e.g.* from *Thermus* sp. (Shaw *et al*, 1995). Moreover, their application is beneficial because such enzymes do not effect dough's rheological properties due to their low activity at moderate temperatures. However, overdosing with thermostable α-amylases leads to undesirably excessive levels of dextrins caused by delayed inactivation during baking. Long operating times for enzymes at baking temperatures is desirable during the production of pumpernickel. That type of bread, prepared from rye flour, has a sweet taste caused by the sugars accumulated during prolonged baking up to 20-24 h at a temperature that does not result in inactivation of the amylolytic enzymes.

In the baking industry glucoamylases are also used to assist the conversion of starch into fermentable sugars. They are especially necessary to improve bread crust colour which is the result of Maillard reactions and intensified by released glucose. Glucoamylases are also added in combination with fungal α-amylases to chilled or frozen dough because they ensure the presence of sufficient quantities of fermentable sugars for yeast when it is time for baking.

Amylolytic enzymes are also used for the formation of the low-molecular weight carbohydrates utilized by yeast growing on starch-containing materials during brewing or the production of alcoholic drinks. In the traditional brewing processes the α- and β-amylases as well as proteinases originate from barley grains germinated for a period of about seven days. This is followed by a process of kilning in which the grain is heated in order to develop colour and flavour. During the mashing stage the enzymes degrade the starch and proteins present in the malt and the additives prepared from crushed starchy cereals such as maize, sorghum, rice or barley. The mixture is then filtered and the clear liquid is boiled in order to inactivate the enzymes. The products of the enzymatic degradation of the malt and additives *i.e.* simple sugars, amino acids and oligopeptides are utilized by yeast, *e.g. Saccharomyces cerevisiae* for the production of alcohol and carbon dioxide, new yeast cells and flavouring components.

Considerable savings can be achieved by replacing some of the malt by unmalted cereals and commercial α-amylases, β-amylases, glucoamylases, β-glucanases and proteinases. This enables better control of the process because the content and activity of the enzymes in the malt are highly variable. Native starch from cereals is resistant to enzyme action and needs to be gelatinized at an elevated temperature. However, gelatinized cereals are very viscous and difficult to handle. Application of thermostable α-amylase at this stage of the process prevents an undesirable increase in viscosity.

Immobilized glucoamylases are used in the modern technologies of low-calorie beer production. Under traditional brewing conditions a large amount of starch is converted into non-fermentable dextrins which are carried through to the final product. Passing the fermenting beer through a reactor containing immobilized glucosidase leads to the break-down of these dextrins to glucose which is then almost completely transformed into alcohol. Additionally, none of enzyme contaminates the final product.

REFERENCES

Ara, K., Saeki, K., Igarashi, K., Takaiwa, M., Uemura, T., Hagihara., H., Kawai, S. and Ito, S. (1995) Purification of an alkaline amylopullulanase with both α-1,4 and α-1,6 hydrolytic activity from alkalophilic *Bacillus* sp. Biochim. Biophys. Acta, *1243*, 315–324.

Christophersen, C. and Otzen, D.E. (1998) Enzymatic characterization of Novamyl, a thermostable α-amylase. Starch/Stärke, *50*, 39–45.

Crabb, D.W. and Mitchinson, C. (1997) Enzymes involved in the processing of starch to sugars. TIBTECH, *15*, 349–353.

Di Lernia, I., Morana, A., Ottombrino, A., Fusco, S., Rossi, M. and De Rosa, M. (1998) Enzymes from *Sulfolobus shibatae* for the production of trehalose and glucose from starch. Extremophiles, *2*, 409–416.

Dong, G., Vieille, C., Savchenko, A. and Zeikus, J.G. (1997) Cloning, sequencing and expression of the gene encoding extracellular α-amylase from *Pyrococcus furiosus* and characterisation of the recombinant enzyme. Appl. Envir. Microbiol. *63*, 3569–3576.

Frillingos, S., Linden, A., Niehaus, F., Vargas, C., Nito, J.J., Ventosa, A., Antranikian, G. and Drainas, C. (2000) Cloning and expression of α-amylase from the hyperthermophilic archaeon *Pyrococcus woesei* in the moderately halophilic bacterium *Halomonas elongata*. J. Appl. Microbiol. *88*, 495–503.

Grzybowska, B., Szweda, P. and Synowiecki, J. (2004) Cloning of the thermostable α-amylase gene from *Pyrococcus woesei* in *Escherichia coli*. Isolation and some properties of the enzyme. Molecular Biotechnol. *26*, 101–109.

Guzman-Maldonado, H. and Paredes-Lopez O. (1995) Amylolytic enzymes and products derived from starch. Crit. Rev. Food Sci. Nutr. *35*, 373–403.

Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. Biochem. J. *316*, 695–696.

James, J.A. and Lee, B.H. (1997) Glucoamylases: microbial sources, industrial applications and molecular biology-a review. J. Food Biochem. *21*, 1–52.

Janeček, S. (1997) α-Amylase family: molecular biology and evolution. Progr. Biophys Molec. Biol. *67*, 67–97.

Kelly, C.T. and Fogarty, W.M. (1983) Microbial α-glucosidases. Process Biochem. *18*, 6–12.

Kim, C.H., Nashiru, O. and Ko, J.H. (1996) Purification and biochemical characterization of pullulanase from *Thermus caldophilus*. FEMS Microbiol. Lett. *138*, 147–152.

Koch, R., Spreinant, A., Lemke, K. and Antranikian, G. (1991) Purification and properties of hyperthermoactive α-amylase from the archaebacterium *Pyrococcus woesei*. Arch. Microbiol. *155*, 572–578.

Laderman, K.A., Davis, B.R., Krutzsch, H.C., Lewis, M.S., Griko, Y.V., Privalov, P.L. and Anfinsen, C.B. (1993) The purification and characterization of an extremely thermostable α-amylase from the hyperthermophilic archaebacterium *Pyrococcus furiosus*. J. Biol. Chem. *268*, 24394–24401.

Lai, J.T., Wu, S.C. and Liu, H.S. (1998) Investigation on the immobilization of *Pseudomonas* isoamylase onto polysaccharide matrices. Bioprocess Eng. *18*, 155–161.

Lama, L., Nicolaus, B., Trincone, A., Morzillo, P., De Rosa, M. and Gambacorta, A. (1990) Starch conversion with immobilized thermophilic archaebacterium *Sulfolobus solfataricus*. Biotechnol. Lett. *12*, 431–432.

Legin, E., Copinet, A., Duchiron, F. (1998) A single step high temperature hydrolysis of wheat starch. Starch/Stärke, *50*, 84–89.

Leuschner, C. and Antranikian, G. (1995) Heat-stable enzymes from extremely thermophilic and hyperthermophilic micro-organisms. World J. Microbiol. Biotech. *11*, 95–114.

Machius, M., Vertesy, L., Huber, R. and Wiegand, G. (1996) Carbohydrate and protein-based inhibitors of porcine pancreatic alpha-amylase: structure analysis and comparison of their binding characteristics. J Mol Biol *260*, 409–421.

Martin, M.L. and Hoseney, R.C. (1991) A mechanism of bread firming. II role of starch hydrolysing enzymes. Cereal Chem. *68*, 503–509.

MacGregor, E.A., Janeèek, S. and Svensson, B. (2001) Relationship of sequence and structure to specificity in the α-amylase family enzymes. Biochim. Biophys. Acta, *1546*, 1–20.

Nakada, T., Ikegami, S., Chaen, H., Kubota, M., Fukuda, S., Sugimoto, T., Kurimoto, M. and Tsujisaka, Y. (1996) Purification and characterization of thermostable maltooligosyl trehalose synthase from the thermoacidifilic archaebacterium *Sulfolobus acidocaldarius*. Biosci. Biotech. Biochem. *60*, 263–266.

Nielsen, J.E. and Borchert, T.B. (2000) Protein engineering of bacterial α-amylases. Biochim. Biophys. Acta, *1543*, 253–274.

Nishimoto, T., Nakada, T., Chaen, H., Fukuda, S., Sugimoto, T., Kurimoto, M. and Tsujisaka,Y. (1996) Purification and characterization of a thermostable trehalose synthase from *Thermus aquaticus*. Biosci. Biotech. Biochem. *60*, 835–839.

Richards, A.B., Krakowska, S., Dexter, L.B., Schmid, H., Wolterbeek, A.P.M., Waalkens-Berendsen, D.H., Arai, S. and Kurimoto, M. (2002) Trehalose: a review of properties, history of use and human tolerance, and results of multiple safety studies. Food Chem. Toxicol. *40*, 871–898.

Shaw, J.F., Lin, F.P. and Chen, H.C. (1995) Purification and properties of an extracellular α-amylase from *Thermus* sp. Bot. Bull. Acad. Sin. *36*, 195–200.

Shaw, P.E., Tatum, J.H. and Wilson, C.W. (1984) Improved flavour of novel orange and grapefruit juices by removal of bitter components with β-cyclodextrin polymer. Cereal Foods World, *33*, 929–934.

Shen, G.J., Saha, B.C., Lee, Y.E., Bhatnagar, L. and Zeikus, J.G. (1988) Purification and characterization of a novel thermostable β-amylase from *Clostridium thermosulphurogenes*. Biochem J., *254*, 835–840.

Szejtli, J. (1982) Cyclodextrins in food cosmetics, and toiletries. Starch, *11*, 379–390.

Synowiecki, J., Sikorski, Z.E. and Naczk, M. (1982) Immobilisation of amylases on krill chitin. Food Chem. *8*, 239–246.

Vieille, C. and Zeikus, G.J. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms of thermostability. Microbiol. Mol. Biol.Rev. *65*, 1–43.

Zdzieblo, A. and Synowiecki, J. (2002) New source of the thermostable α-glucosidase suitable for single step starch processing. Food Chem. *79*, 485–491.

CHAPTER 3

CELLULASES FOR BIOMASS CONVERSION

QI XU, WILLIAM S. ADNEY, SHI-YOU DING AND MICHAEL E. HIMMEL*
National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO, USA
**mike_himmel@nrel.gov*

1. INTRODUCTION

A primary goal of the National Energy Policy is to increase United States energy supplies using a more diverse mix of domestic resources and to reduce our dependence on imported oil. In 2002, fossil fuels, which are finite and non-renewable, supplied 86% of the energy consumed in this country. Even more alarming is that the United States imports over half (62%) of its petroleum, and dependency is increasing. In particular, gasoline and diesel constituted 98% of domestic transportation motor fuels in 2004. The United States gasoline consumption alone was about 138 billion gal/year in 2004. Corn ethanol supplies most of the remaining 2%. Bioethanol from cornstarch provides around 3 to 4 billion gallons of oxygenate that is splash-blended with gasoline to produce the common "gasohol." An Oak Ridge National Laboratory study, published in April 2005, indicated a potentially renewable feedstock base in the United States of over a billion tons per year that could generate 30% of current petroleum consumption. The feedstocks included forest thinnings, crop residues, bioenergy crops and wastes. Achieving this increase will require substantial RandD in feedstock production, harvesting, and land use. In order to efficiently utilize these lignocellulosic feedstocks, powerful new plant cell wall degrading enzymes will be required, especially cellulases.

Feedstock costs will be a major component of the commodity end-product price. Therefore, yield of lignocellulose-derived sugars is perhaps of highest priority. Another impact on feedstock yield is associated with cellulases and other polysaccharide-degrading enzymes. These enzyme preparations must work efficiently to convert the dominant polysaccharides to monomers. Currently, high loadings of cellulases are needed to reach 95% conversion of cellulose in pretreated biomass in 3–5 days in a simultaneous saccharification and fermentation (SSF)

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 35–50

experiment *i.e.* 2.2 lb (1 kg) of cellulase for 110 lb (50 kg) of cellulose (Grohmann *et al.*, 1991). Cellulase preparations are expensive in the biorefinery context for two reasons: (1) the enzyme source, usually *Trichoderma reesei*, is costly to grow and induce and has limited cellulase productivity; and (2) specific enzyme performance or activity has not been improved by discovery or protein engineering in 30 years of research. A consequence of recent work announced by Genencor International was the significant breakthrough in reducing the cost to produce *T. reesei* cellulases from about $5/gal of ethanol to around $0.20/gal (Mitchinson *et al.*, 2005).

The biomass feedstocks most commonly considered for conversion are agricultural wastes, energy crops (perennial grasses and trees), and forest waste. The fermentable fractions of these feedstocks include cellulose (β-1,4-linked glucose) and hemicellulose, a substantial heterogeneous fraction composed of xylose and minor five- and six-carbon sugars. Although it is an abundant biopolymer, cellulose is unique because it is highly crystalline, water insoluble, and highly resistant to depolymerization. The definitive enzymatic degradation of cellulose to glucose, probably the most desirable fermentation feedstock, is generally accomplished by the synergistic action of three distinct classes of enzymes:

i) The "endo-1, 4-β-glucanases" or 1,4-β-D-glucan 4-glucanohydrolases (EC 3.2.1.4), which act randomly on soluble and insoluble 1,4-β-glucan substrates and are commonly measured by detecting the reducing groups released from carboxymethylcellulose (CMC
ii) The "exo-1,4-β-D-glucanases," including both the 1,4-β-D-glucan glucohydrolases (EC 3.2.1.74), which liberate D-glucose from 1,4-β-D-glucans and hydrolyze D-cellobiose slowly, and 1,4-β-D-glucan cellobiohydrolase (EC 3.2.1.91), which liberates D-cellobiose from 1,4-β-glucans.
iii) The "β-D-glucosidases" or β-D-glucoside glucohydrolases (EC 3.2.1.21), which act to release D-glucose units from cellobiose and soluble cellodextrins, as well as an array of glycosides.

2. HISTORICAL MODELS FOR CELLULASE ACTION

From the published work of de Bary in the 18th Century (de Bary, 1886), scientists were aware that an enzyme (from fungal extracts) degraded plant cell-wall polysaccharides. In 1890, Brown and Morris (1890) concluded that the cellulose-dissolving power in Barley extracts is due to a special enzyme and that this enzyme is not diastase (the name for starch degrading enzymes at the time). Newcombe (1899) showed conclusively that the cellulose-degrading enzyme (named cytase or cytohydrolyst) in Barley malt was distinct from starch degrading enzymes. Interestingly, the German literature at the time referred to cellulose degrading enzymes as "celluloselosendes enzyms", or cellulose-loosening enzymes (Reinitzer, 1897). From our review of the literature, the first reference to "cellulases" as enzymes that degrade cellulose was made by Pringsheim (1912). By the 1920s, evidence was mounting that these enzymes were actually proteins and that proteins were discrete chemical entities. However, the answer to this question

had to wait for sufficiently sophisticated protein purification techniques to be developed.

The search for biological causes of cellulose hydrolysis did not begin in earnest until World War II. The U.S. Army mounted a basic research program to understand the causes of deterioration of military clothing and equipment in the jungles of the South Pacific – problem that was wreaking havoc with cargo shipments during the war. Out of this effort to screen thousands of samples collected from the jungle came the identification of what has become one of the most important organisms in the development of cellulase enzymes – *Trichoderma viride* (eventually renamed *Trichoderma reesei*). In 1973, the army was beginning to look at cellulases as a means of converting solid waste into food and energy products (Brandt *et al.*, 1973). By 1979, genetic enhancement of *T. reesei* had already produced mutant strains with up to 20 times the productivity of the original organisms isolated from New Guinea (Mandels *et al*, 1971; Montenecourt and Eveleigh, 1979). For roughly 20 years, cellulases made from submerged culture fungal fermentations have been commercially available. In another ironic twist, the most lucrative market for cellulases today is in the textile industry where "partial system" preparations displaying minimal cellulose degradation are employed.

In many ways, however, our understanding of cellulases is in its infancy compared to other enzymes. There are some good reasons for this. Cellulase–cellulose systems involve soluble enzymes working on insoluble substrates. The jump in complexity from homogeneous enzyme-substrate systems is tremendous. It became clear fairly quickly that the enzyme known as "cellulase" was really a complex system of enzymes that work together synergistically to attack native cellulose. Early views of cellulase action considered the system to embody a C_1 activity, which acts in an unspecified way to disrupt the crystalline structure of cellulose, and the C_x activity, which encompasses all β-1,4-glucanase action, including the exoglucanases and the endoglucanases (Reese *et al*, 1950). Thus, the picture of the cellulase system from the view of the late 1960s was limited by proposition of the as-yet-uncharacterized C_1 factor (King and Vessal, 1969). During the next decade, the fungal cellulase system was interpreted largely in terms of substantial biochemical and molecular biological developments in the *Trichoderma reesei* system. In many ways, this system was the developmental archetype cellulase system. Many reviews have adequately described the 20-plus years of systematic research conducted at the Army Natick Laboratory on this subject (Mandels and Reese, 1964).

3. NON-CELLULOSOMAL CELLULASES (THE FUNGAL MODEL)

The cellulase field moved ahead dramatically in the late 1980s when Abuja and co-workers reported the tertiary structure of *T. reesei* CBH I and CBH II (Abuja *et al.*, 1988a,b). This structure, determined by small-angle X-ray scattering (SAXS) data, depicted these proteins as two domain proteins whose form resembles tadpoles. This now-familiar structure is composed of a large core (catalytic) domain; a small cellulose-binding domain (CBD); and a linker, or hinge, peptide connecting the

two. In the case of *T. reesei* CBH II, the core protein itself has been shown to cause disruption in cellulose microfibril structure (Woodward *et al*., 1992). The core domains of *T. reesei* CBH I and CBH II have now been shown to possess seven and four active site glucopyranoside "subsites," respectively (Teeri *et al*., 1994). Furthermore, CBH I produces hydrolysis products with a retained stereochemistry at the anomeric carbon, while CBH II causes an inversion of the anomeric hydroxyl to the α-form. The cartoon shown in Fig. 1 depicts an idealized cellulase enzyme, based on the general shapes and orientation of the catalytic and cellulose binding domains. In the *T. reesei* enzymes (and in many other cellulases), the linker peptide is a highly glycosylated region unusually rich in serine, threonine, and proline amino acid residues. This linker region is also the site of proteolytic cleavage accomplished by several general serine proteases. Interestingly, there appears to be a considerable level of conservation in nature for this general structure, as evidenced by homologies in the linker peptide found for an *Aspergillus niger* protease, an α-amylase from *Hordeum vulgare*, and a α-amylase from *Saccharomyces diastaticus* (Claeyssens *et al.*, 1990).

Elucidation of the structure of the 36 amino acid peptide Type 1 CBD from *T. reesei* CBH I by C^{13} NMR in 1989 revealed the presence of a strongly hydrophobic peptide "face" (Kraulis *et al*., 1989). Today, most workers in the field conclude that the CBM (the CBD was renamed, carbohydrate binding module) plays a role in stabilizing cellulase attachment to the cellulosic surface. The cartoon shown in Fig. 2 represents the surface-binding configuration of a cellobiohydrolase, a general mechanism for CBM/cellulose interaction suggested by Rouvinen and coworkers (1990). These results strongly support the idea of well-ordered hydrophobic interaction with the surface of the cellulose at the CBM.

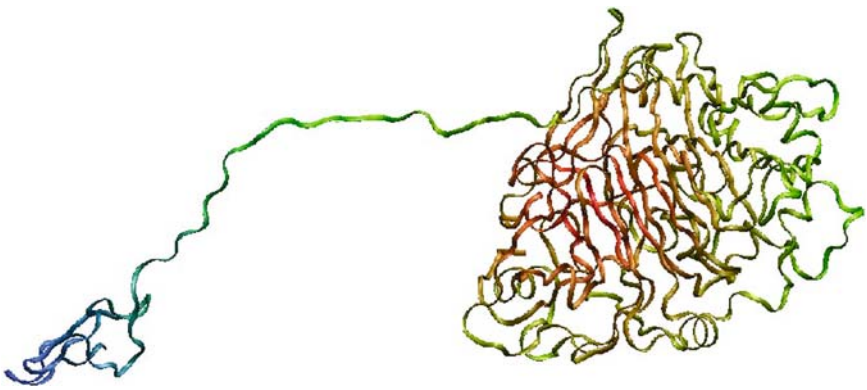


Figure 1. The proposed structure of *T. reesei* CBH I showing the cellulose binding domain (CBM), a 26-amino acid linker peptide, and catalytic domain. The catalytic domain of CBH I contains a 10 subsite active site tunnel from which cellobiose is released as the end product

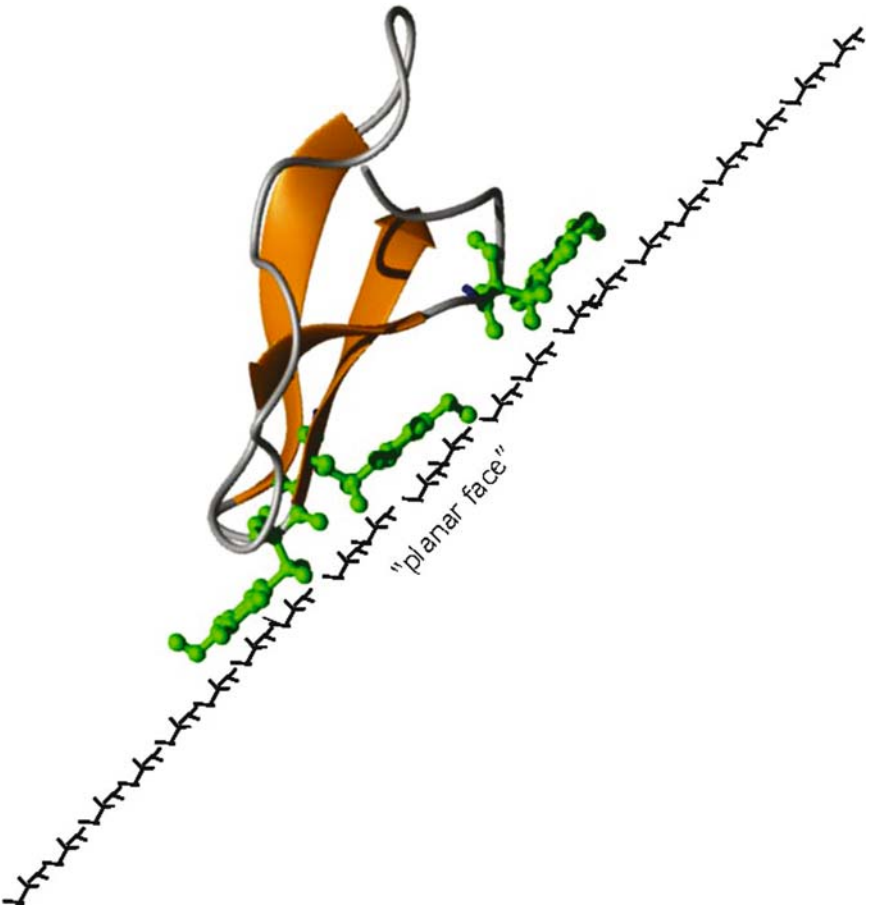


Figure 2. Depiction of a type I CBM from *T. reesei* interacting with the 1,0,0 or planar face of cellulose. This family of CBMs is distinguished by small domains all containing three tyr residues placed in nearly a co-linear pattern on the cellulose interaction surface

4. CELLULASE SYNERGISM

An early work by Gilligan and Reese (1954) showed that the amount of reducing sugar released from cellulose by the combined fractions of fungal culture filtrate was greater than the sum of the amounts released by the individual fractions. Since that time, many investigators have used a variety of fungal preparations to demonstrate a synergistic interaction between homologous exo- and endo-acting cellulase components (Li *et al*., 1965; Selby, 1969; Wood, 1969; Halliwell and Riaz, 1970; Wood and McCrae, 1979; Eriksson, 1975; Petterson, 1975; McHale and Coughlan, 1980). Cross-synergism between endo- and exo-acting enzymes from filtrates of different aerobic fungi has also been demonstrated several times (Selby, 1969; Wood, 1969; Wood and McRae, 1977; Coughlan *et al*., 1987).

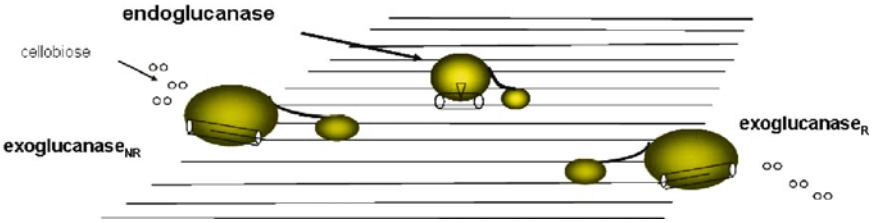


Figure 3. Concept of endo-exo and exo-exo cellulase synergism thought to play a key role in the function of both fungal and bacterial cellulases. In general, the endoglucanases produce "nicks" in the cellulose strands and these free ends are targeted by the exoglucanases (cellobiohydrolases in fungi)

Fägerstam and Pettersson (1980) first reported exo-exo synergism in 1980. The concepts of exo-endo and exo-exo synergism are shown diagrammatically in Fig. 3. As shown in this drawing, exo-endo synergism is explained best in terms of providing new sites of attack for the exoglucanases. These enzymes normally find available cellodextrin "ends" at the reducing and non-reducing termini of cellulose microfibrils. Random internal cleavage of surface cellulose chains by endoglucanases provides numerous additional sites for attack by cellobiohydrolases. Therefore, each hydrolytic event by an endoglucanase yields both a new reducing and a new non-reducing site. Thus, logical consideration of catalyst efficiency dictates the presence of exoglucanases specific for reducing termini and non-reducing termini. Indeed, an X-ray crystallography study reported by Teeri *et al.* (1994) confirmed that the reducing terminus of a cellodextrin can be shown in proximal orientation to the active site tunnel; *i.e.* reducing end in first, of *T. reesei* CBH I. Earlier kinetic data had already confirmed that *T. reesei* CBH II preferred the non-reducing approach to the cellulose chain (Claeyssens *et al.*, 1989).

Synergism between fungal and bacterial exo- and endo-acting components was first proposed by Eveleigh (1987) and reported by Wood (1988). These observations have most recently been extended by Irwin *et al.* (1993) and in the authors' laboratory (Baker *et al.*, 1998). This principle of interspecific interchangeability of cellulase components is now the cornerstone of recombinant cellulase system design and construction. If indeed cellulase component enzymes are truly generalized in both structure and function, components may be selected and combined from a wide array of source organisms to form novel enzyme cocktails. For example, *T. reesei* CBH I has been shown to be a powerful element in multi-enzyme mixtures using either fungal or bacterial endoglucanases.

5. MODERN CLASSIFICATION OF CELLULASES

Today, more than 90 families of glycosyl hydrolases have been identified (Carbohydrate-Active Enzymes server at URL: *http://afmb.cnrs-mrs.fr/CAZY/*; Coutinho and Henrissat, 1999). This classification system provides a powerful tool for glycosyl hydrolase enzyme engineering studies, because many enzymes critical

for industrial processes have not yet been crystallized or subjected to structure analysis. Glycosyl hydrolase (GH) families harboring enzymes known to play a role in cellulose degradation are Families 1, 3, 5, 6, 7, 9, 10, 12, 16, 44, 45, 48, 51, 61 and 74. These cellulase enzymes have been grouped using protein sequence alignment algorithms Hydrophobic Cluster Analysis. The cellulase Families include members from widely different fold types, *i.e.* the TIM-barrel, β/α-barrel variant (a TIM-barrel-like structure that is imperfectly superimposable on the TIM-barrel template), β-sandwich, and α-helix circular array. This diversity in cellulase fold structure must be taken into account when considering the transfer and application of design strategies between different cellulases.

Protein domains are grouped into four general structural categories (all-alpha, all-beta, alpha + beta, and alpha/beta) (Levitt and Clothia, 1976; Efimov, 1994). Proteins of the all-alpha class are usually comprised of multiple alpha helices which may be oriented along a common bundle axis or oriented randomly (Harris *et al.*, 1994). Proteins of the all-beta class contain beta-strands which can be oriented either parallel, or antiparallel, or a mixture of the two. The beta + alpha and alpha/beta categories are distinguished by considering that alpha/beta proteins have alternating beta-strand and alpha-helical segments, whereas alpha + beta proteins tend to contain regions definable as "mostly alpha" and "mostly beta" (Orengo and Thornton, 1993). A common example of the alpha/beta class is the TIM-barrel, named after the archetype of this fold, triose phosphate isomerase. In TIM-barrel proteins, the internal barrel is comprised of 8 parallel beta-strands, while the outer shell contains 8 alpha helices oriented with a cant relative to the axis of the barrel. Some protein domains do not fall into one of these categories and are grouped as irregular folds. Proteins representative of these domain (or fold) classes are myoglobin (all-alpha helix), immunoglobulin (all-beta strand), cytochrome b5 (alpha + beta), and triose phosphate isomerase (alpha/beta) (Levitt and Clothia, 1976). It is inferred that all proteins, which have recognizable sequence similarity, will have the same fold type. In many cases, the fold will be unique to that single family of proteins and such folds are known as structural singlets (Orengo and Thornton, 1993). In other cases, a domain structures (fold) may be shared by two or more proteins that appear unrelated by sequence and function. Such folds have been termed superfolds

Cellulases are generally defined as enzymes which hydrolyze the β-1,4-glucosidic bonds within the chains that comprise the cellulose polymer. A narrower definition of "true cellulase" has also been used, which are enzymes can act alone on insoluble cellulose. Tables 1 and 2 show the major families of cellulases described in the Carbohydrate-Active Enzymes (CAZy) server (http://afmb.cnrs-mrs.fr/CAZY/). As it is indicated in Table 2 there are two major catalytic mechanisms which lead to either retention or inversion of the configuration of the anomeric hydroxyl. In all cases, the proton donor and nucleophile/base are Glu or Asp. Most cellulase families are distributed among bacteria, fungi, and plants. Interestingly, GH7 and GH61 are found only in fungi, whereas family GH44 is found only in bacteria. The GH7 (cellobiohydrolase) is the most active exoglucanase known, and is widely believed to act

Table 1. EC Numbers And Cellulase Families

E.C. #	Reaction	Other Names	Family
E.C.3.2.1.4 Cellulase	*Endohydrolysis of 1,4-beta-D-glucosidic linkages in cellulose, lichenin and cereal beta-D-glucans.*	*Endoglucanase. Endo-1,4-beta-glucanase. Carboxymethyl cellulase. Endo-1,4-beta-D-glucanase. Beta-1,4-glucanase. Beta-1,4-endoglucan hydrolase. Celludextrinase. Avicelase.*	5, 6, 7, 8, 10, 12, 44, 45, 48, 51, 61, 74
E.C.3.2.1.6 Endo-1,3(4)-beta-glucanase.	*Endohydrolysis of 1,3- or 1,4-linkages in beta-D-glucans when the glucose residue whose reducing group is involved in the linkage to be hydrolyzed is itself substituted at C-3.*	*Endo-1,4-beta-glucanase. Endo-1,3-beta-glucanase. Laminarinase.*	16
E.C.3.2.1.21 Beta-glucosidase.	*Hydrolysis of terminal, non-reducing beta-D-glucose residues with release of beta-D-glucose.*	*Gentobiase. Cellobiase. Amygdalase.*	1, 3, 9
E.C.3.2.1.91 Cellulose 1,4-beta-cellobiosidase	*Hydrolysis of 1,4-beta-D-glucosidic linkages in cellulose and cellotetraose, releasing cellobiose from the non-reducing ends of the chains.*	*Exoglucanase. Exocellobiohydrolase. 1,4-beta-cellobiohydrolase.*	5, 6, 7, 9, 10, 48,

processively on a single cellulose chain. As is the case for the GH7 family, GH6 contains both endo- and exo-glucanases, however, the exoglucanases in GH6 are found both in bacteria and fungi. GH6 and GH7 exoglucanases act from the non-reducing and reducing termini, respectively. Bacterial exoglucanases are also resident in GH9, GH48 and GH74 and these enzymes are thought to act non-precessively.

6. CELLULOSOMAL CELLULASES

The cellulosome shown in Fig. 4 is an extracellular, multi-protein complex that is produced by a wide range of cellulolytic micro-organisms. It is believed to have the feature of "collecting" and "positioning" cellulose degrading enzymes onto a substrate (Bayer *et al*, 1994). The functional unit of the cellulosome is the "scaffoldin," which is a non-catalytic protein containing repetitive domains (cohesins) for specific interaction with other protein domains, called dockerins. Cellulosomal enzymes contain both a catalytic domain and a binding domain (dockerin). The cellulosome then self-assembles by type-specific recognition of

Table 2. Cellulase Families, Structure, Activity, and Distribution

GH	Structure	Activity	Catalytic Mechanism	Nucleophile /Base	Proton Donor	Bacteria	Fungi	Plant
1	$(\beta/\alpha)_8$	β−glucosidase	Retaining	Glu	Glu	+	+	+
3		β−glucosidase	Retaining	Asp	Glu	+	+	+
5	$(\beta/\alpha)_8$	Endoglucanase,	Retaining	Glu	Glu	+	+	+
6		Endoglucanase, cellobiohydrolase	Inverting	Asp	Asp	+	+	
7	β−jelly roll	Endoglucanase, cellobiohydrolase	Retaining	Glu	Glu		+	
9	$(\alpha/\alpha)_6$	Endoglucanase, cellobiohydrolase	Inverting	Asp	Glu	+	+	+
10	$(\beta/\alpha)_8$	Cellobiohydrolase	Retaining	Glu	Glu	+	+	+
12	β−jelly roll	Endoglucanase	Retaining	Glu	Glu	+	+	
16	β−jelly roll	Endo-1,3(4)-β-glucanase	Retaining	Glu	Glu	+	+	+
44		Endoglucanase	Inverting	N/A	N/A	+		
45		Endoglucanase	Inverting	Asp	Asp	+	+	
48	$(\alpha/\alpha)_6$	Endoglucanase, cellobiohydrolase	Inverting	N/A	Glu	+	+	
51	$(\beta/\alpha)_8$	Endoglucanase	Retaining	Glu	Glu	+	+	+
61	N/A	Endoglucanase	N/A	N/A	N/A		+	
74	7-fold β-propeller	Endoglucanase, cellobiohydrolase	Inverting	Asp	Asp	+	+	

cohesin/dockerin pairs. The scaffoldins can also contain the carbohydrate-binding module (CBM) which serves as an attachment device for harnessing the cellulosome to the cell surface and/or for its targeting to substrate.

6.1. Non-Catalytic Subunit: Scaffoldin

The cellulosome is one of the best-studied protein complexes known to form self-assembled extracellular scaffolds (Bayer *et al*., 2004). The molecular mass of the cellulosome complex was determined to be several MDa. Two types of subunits have been identified from the bacterial cellulosome complex. *Non-catalytic subunits*, called "scaffoldins", serve to position and organize the *enzymatic subunits* and to attach the cellulosome to the cell surface and/or to the substrate – *i.e.* plant cell wall polysaccharides.

The scaffoldins contain multiple copies of *cohesins*, which interact with *dockerin* domains of the enzymatic subunits to form the cellulosome assembly. The cohesins are about 140 amino acids in length and highly conserved in sequence and domain structure. The dockerin domains comprise about 70 amino acids and contain two 22-amino acid duplicated regions, each of which includes an "F-hand" modification of the EF-hand calcium-binding motif. To date, several hundred cohesin and dockerin sequences have been found, mostly from anaerobic bacteria. More than

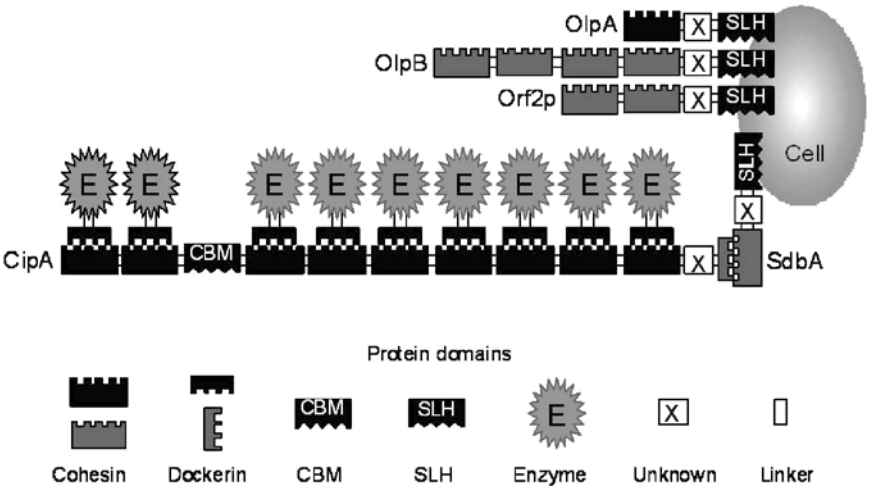


Figure 4. Schematic structure (not scaled) of an example cellulosome complex from *Clostridium thermocellum*. The cellulosome complex are composed of two groups of proteins. One group is non-catalytic proteins (scaffoldin) including CipA, SdbA, Orf2p, OlpA, and OlpB, each of these scaffoldins contain various number of function domains, *i.e.* cohesin domain interacts with same type of dockerin domain (Type-I cohesin-dockerin pair are showing in black and Type-II pair in grey); carbohydrate-binding module (CBM) recognizes polysaccharide substrate; S-layer homologous (SLH) binds to cell surface; and linker between these domains. Another group is catalytic proteins (enzymes), each cellulosomal enzyme contains a Type-I dockerin domain recognizing Type-I cohesin of scaffoldin proteins. In *Clostridium thermocellum*, more than twenty enzymes with various catalytic activities have been identified to be involved in cellulosome complex

a dozen different specificities are currently known which will enable the design and production of numerous types of nano-component systems.

Bacterial cellulosomes are organized by means of a special type of subunit, the scaffoldin, which is comprised of an array of cohesin modules. The cohesin interacts selectively and tenaciously with a complementary type of domain, the dockerin, which is borne by each of the cellulosomal enzyme subunits. The integrity of the complex is thus maintained by the cohesin-dockerin interaction. The first scaffoldin was sequenced from *Clostridium cellulovorans* (Shoseyov *et al.*, 1992). The relationship to the duplicated sequences of cellulosomal enzymes (Salamitou *et al.*, 1992) was later realized when a second scaffoldin, derived from *C. thermocellum*, was sequenced (Gerngross *et al.*, 1993). Today, many scaffoldin genes has been sequenced and characterized from *C. thermocellum, C. josui* (Fujino *et al.*, 1993), *B. cellulosolvens* (Xu *et al.*, 2004a), *A. cellulolyticus* (Xu *et al.*, 2004b), *and R. flavefaciens* (Ding *et al.*, 2001).

The cellulosome system characterized by multiple scaffoldins includes a primary scaffoldin, anchoring scaffoldins, and an "adaptor" scaffoldin. The primary scaffoldin incorporates the enzymatic subunits and usually bears a single CBM domain. The anchoring scaffoldin bears an SLH module for attaching the cellulosome to the cell

wall. The adaptor scaffoldin from *A. cellulolyticus* contains four cohesins and a dockerin, which effectively multiplies the number of enzymes that can be incorporated into the complex. In contrast, the adaptor scaffoldin from *R. flavefaciens* contains a single divergent cohesin and alters the specificity of the primary scaffoldin which expands the repertoire of cellulosomal subunits that can be incorporated into the complex. Scaffoldins have significant diversity in cellulosome architecture, as reflected by the number of cohesins in a given scaffodin and their disposition therein, the presence (or absence) and location of a CBM, and the presence (or absence) of a dockerin and/or SLH module. For example, the *R. flavefaciens* scaffoldin lack an identifiable CBM and SLH, although the cellulosome binds cellulose and is cell associated, an enzyme-bearing CBM might mediate this important function (Rincon *et al*., 2001); the scaffoldin (*scaD*) from *A. cellulolyticus* plays a dual role, both as a primary scaffoldin –capable of direct incorporation of a single dockerin-borne enzyme and as a secondary scaffoldin – one that anchors the major primary scaffoldin, *ScaA*, and its complement of enzymes to the cell surface (Xu *et al.*, 2004b). In the case of mesophilic *Clostridia*, their sacffoldins lack dockerins and conventional SLH domains. However, a similar type of module contained at the *N*-terminus of the *C. cellulovorans* enzyme family-9 enzyme, EngE, has been implicated in mediating cell surface attachment of its cellulosome (Kosugi *et al*., 2002).

6.2. The Cohesin-dockerin Interaction

The first biochemical analyses of the cellulosome complex from *C. thermocellum* indicated an exceptionally strong interaction that rivaled the affinities of the most tenacious biochemical bonds (Lamed and Bayer, 1988; Lamed *et al*., 1983). Subsequent analyses substantiated these claims, and the cohesin-dockerin interaction rates among the most potent protein-protein interactions known in nature (Fierobe *et al.*, 2001; Mechaly *et al*., 2001). The interaction between the two components can be viewed as a kind of plug-and-socket arrangement, whereby the dockerin domain plugs into the cohesin module (Bayer *et al*., 2004).

6.3. Carbohydrate-Binding Modules

Glycosyl hydrolases attach to polysaccharides relatively inefficiently, as their target glycosidic bonds are often inaccessible to the active site of the appropriate enzymes. In order to overcome these problems, many of the glycosyl hydrolases, primarily the noncellulosomal cellulases and related “free” enzymes that hydrolyze insoluble substrates, are modular and comprise catalytic modules appended to one or more non-catalytic CBMs (carbohydrate-binding modules). CBMs primarily promote the association of the enzyme with the substrate (Boraston *et al*., 2004; Bayer *et al*., 2004).

CBMs are divided into families based on amino acid sequence similarity. There are currently 43 defined families and these displayed substantial variation in ligand specificity (see http://afmb.cnrs-mrs.fr/CAZY/CBM.html). Thus there are characterized CBMs that recognize crystalline cellulose, non-crystalline cellulose,

chitin, β-1,3-glucans and β-(1,3)-(1,4) mixed linkage glucans, xylan, mannan, galactan, and starch. Some CBMs display "lectin-like" specificity and bind to a variety of cell-surface glycans (Boraston *et al.*, 2004; Sorimachi *et al.*, 1996, 1997; Williamson *et al.*, 1997; Sigurskjold *et al.*, 1994). Based on structural and functional similarities, CBMs are been grouped into three types:

6.3.1. Type A Surface

Binding CBMs This class of CBMs binds to insoluble, highly crystalline cellulose and/or chitin. The aromatic amino acid residues play key role in the binding sites. The planar architecture of the binding sites is thought to be complementary to the flat surfaces presented by cellulose or chitin crystals (Bayer *et al.*, 1999). The substrate binding site comprises the "hydrophobic" face of cellulose (Bayer *et al.*, 1999). Upon binding to the substrate, the cellulosome is thought to undergo a supramolecular rearrangement so that the components redistribute to interact with the different target substrate. For this purpose, the various cellulosomal enzymes include different types of CBMs from different families that exhibit appropriate specificities that complement the action of the parent enzyme (Bayer *et al.*, 2004).

6.3.2. Type B Polysaccharide-Chain-Binding CBMs

This class of CBMs binds to individual glycan chains. As with type A CBMs, aromatic residues play a pivotal role in ligand binding, and the orientation of these amino acids are key determinants of specificity. The binding sites often described as grooves or clefts, and comprise several sub-sites able to accommodate the individual sugar units of the polymeric ligand (Simpson *et al.*, 2000). In sharp contrast with the Type A CBMs, direct hydrogen bonds also play a key role in the defining the affinity and ligand specificity of Type B glycan chain binders (Notenboom *et al.*, 2001; Xie *et al.*, 2001).

6.3.3. Type C Small-Sugar-Binding CBMs

This class of CBMs has the lectin-like property of binding optimally to mono-, di-, or tri-saccharides and thus lacks the extended binding-site grooves of type B CBMs. The distinction between Type B CBMs and Type C CBMs can be subtle (Boraston *et al.*, 2003).

6.3.4. Type D CBMs

This class of CBMs is always found in close spatial proximity with the catalytic domains of their respective proteins. Examples include the cellulase family 9 enzymes from *T. fusca* (Sakon *et al.*, 1997).

7. OUTLOOK FOR CELLULASE RESEARCH

It is now clear that cutting-edge and efficient biochemical technologies must be used to reduce the cost of cellulase activities delivered to the SSCF bioethanol process. The current estimate for NREL Proven Technologies and Best of Industry

Technologies yields cellulase costs to the bioethanol process of $0.32 and $0.18 per gallon ethanol produced, respectively. These costs must be reduced to less than $0.05 per gallon ethanol by 2020 and this requires further increases in specific activity or production efficiency or some combination thereof (Wooley and Ruth, 1999). It is most likely that the needed further improvements in cellulase performance will come via continued research aimed at understanding the basic principles by which these enzymes function on microcrystalline cellulose surfaces. Specifically, the mode of action of the "processive" enzymes, such as *T. reesei* CBH I and CBH II, must be more deeply understood before further improvement in activity via enzyme engineering tools can be realized.

ACKNOWLEDGEMENT

This work was funded by the U.S. Department of Energy's Office of the Biomass Program.

REFERENCES

Abuja, P.M., Schmuck, M., Pilz, I., Tomme, P., Claeyssens, M. (1988a) Structural and functional domains of cellobiohydrolase I from *T. reesei* - a Small-Angle X-Ray-Scattering Study of the Intact Enzyme and Its Core. Eur. Biophys. J. *15*, 339–342.

Abuja, P.M., Pilz, I., Claeyssens, M. and Tomme, P. (1988b) Domain structure of cellobiohydrolase II as studied by small angle X-ray scattering: close resemblance to cellobiohydrolase I. Biochem. Biophys. Res. Commun. *156*, 180–185.

Baker, J.O., Ehrman, C.I., Adney, W.S., Thomas, S.R., Himmel, M.E. (1998) Hydrolysis of cellulose using ternary mixtures of purified cellulases. Appl. Biochem. Biotechnol. *70/72*, 395–403.

de Bary, A. (1886) Ueber einige sclerotinien und sclerotienkrankheiten. Bot. Zeit. XLIV: 377.

Bayer, E.A., Morag, E. and Lamed, R. (1994) The cellulosome — a treasure-trove for biotechnology. Trends in Biotechnology *12*, 378–386.

Bayer, E.A., Ding, S.-Y., Mechaly, A., Shoham, Y. and Lamed, R. (1999) Emerging phylogenetics of cellulosome structure, In Recent Advances In Carbohydrate Bioengineering, Gilbert, H.J., Davies, G, Henrissat, B., Svensson, B. (Eds), The Royal Society of Chemistry, Cambridge, p. 189–201.

Bayer, E.A., Belaich, J.P., Shoham, Y. and Lamed, R. (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annual Review of Microbiology *58*, 521–554.

Boraston, A.B., Notenboom, V., Warren, R.A., Kilburn, D.G., Rose, D. R. and Davies, G. (2003) Structure and ligand binding of carbohydrate-binding module csCBM6-3 reveals similarities with fucose-specific lectins and "galactose-binding" domains. J. Mol. Biol. 327(3):659–669.

Boraston, A.B., Bolam, D.N., Gilbert, H.J. and Davies, G.J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. *382*, 769–781.

Brandt, D., Hontz, L. and Mandels, M. (1973) Engineering aspects of the enzymatic conversion of waste cellulose to glucose. AIChE Symposium Series *69*, 127.

Brown, H.T., and Morris, G.H. (1890). Researches on the germination of some of the Graminae. J. Chem. Soc. Lond. *57*, 458–528.

Claeyssens, M., Tilbeurgh, H.V., Tomme, P., Wood, T.M. and McCrae, S. I. (1989) Fungal Cellulase Systems: Comparison of the specificities of the cellobiohydrolases isolated from *Penicillium pinophilum* and *Trichoderma reesei*. Biochem. J. *261*, 819–826.

Claeyssens, M. and Tomme, P. (1990) In Kubicek, C. P., Eveleigh, D. E., Esterbauer, H., Steiner, W. and Kubicek-Pranz, E.M. (Eds.), *Trichoderma reesei* Cellulases: Biochemistry, Genetics, Physiology and Application, London, the Royal Society of Chemistry, pp. 1–11.

Coughlan, M.P., Moloney, A.P., McCrae, S.I. and Wood, T.M. (1987) Cross-synergistic interactions between components of the cellulase systems of *Talaromyces emersonii*, *Fusarium solani*, *Penicillium funiculosum*, and *Trichoderma koningii*. Biochem. Soc. Transactions *15*, 263–264.

Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. In Recent Advances in Carbohydrate Bioengineering, Gilbert, H.J., Davies, G., Henrissat, B. and B. Svensson (eds.), The Royal Society of Chemistry, Cambridge, pp. 3–12.

Ding, S.-Y., Rincon, M.T., Lamed, R., Martin, J.C., McCrae, S.I., Aurilia, V., Shoham, Y., Bayer, E.A. and Flint, H.J. (2001) Cellulosomal scaffoldin-like proteins from *Ruminococcus flavefaciens*. J. Bacteriol. *183*, 1945–1953.

Efimov, A.V. (1994) Favoured structural motifs in globular proteins. Structure *2*, 999–1002.

Eriksson, K.E. (1975) Enzyme mechanisms involved in the degradation of wood components. In Bailey, M., Enari T.M. and Linko, M. (Eds.), Symposium on Enzymatic Hydrolysis of Cellulose. Helsinki, SITRA, p. 263.

Eveleigh, D.E. (1987) Cellulases: A perspective. Phil. Trans. Roy. Soc. Lond., A321:435–447.

Fägerstam, L. G. and Pettersson, L. G. (1980) The 1,4-β-glucan cellobiohydrolases of *T. reesei* QM 9414. A new type of cellulolytic synergism. FEBS Lett. *119*, 97.

Fierobe, H.P., Mechaly, A., Tardif, C., Belaich, A., Lamed, R., Shoham, Y., Belaich, J.P. and Bayer, E.A. (2001) Design and production of active cellulosome chimeras: selective incorporation of dockerin-containing enzymes into defined functional complexes. J. Biol. Chem. *276*, 21257–21261.

Fujino, T., Karita, S. and Ohmiya, K. (1993) Nucleotide-sequences of the celB gene encoding endo-1,4-beta-glucanase-2, Orf1 and Orf2 forming a putative cellulase gene-cluster of *Clostridium josui*. J. Ferment. and Bioeng. 76(4):243–250.

Gerngross, U.T., Romaniec, M.P. M., Kobayashi, T., Huskisson, N.S. and Demain, A.L. (1993) Sequencing of a *Clostridium thermocellum* gene (cipA) encoding the cellulosomal S(L)-protein reveals an unusual degree of internal homology. Mol. Microbiol. 8(2):325–334.

Gilligan, W. and Reese, E.T. (1954) Evidence for multiple components in microbial cellulases. Can. J. Microbiol. *1*, 90.

Grohmann, K., Wyman, C.E. and Himmel, M.E. (1991) Potential for fuels from biomass and wastes. In Rowell, R., Narayan, R. and Schultz, T. (eds), Emerging Materials and Chemicals from Biomass. ACS Series *476*, 354–392.

Halliwell, G. and Riaz, M. (1970) The formation of short fibres from native cellulose by components of Trichoderma koningii cellulase. Biochem. J. *116*, 35–42.

Harris, N.L., Presnell, S.R. and Cohen, F.E. (1994) Four helix bundle diversity in globular proteins. J. Mol. Biol. *236*, 1356–1368.

Irwin, D.C., Spezio, M., Walker, L.P. and Wilson, D.B. (1993) Activity studies of eight purified cellulases: specificity, synergism, and binding domain effects. Biotech. Bioeng. *42*, 1002–1013.

King, K.W. and Vessal M.I. (1969) Enzymes of the cellulase complex. In Hajny, G. J. And Reese E. T. (eds), Cellulases And Their Applications. Advances in Chemistry Series *95*, 7–25.

Kosugi, A., Murashima, K., Tamaru, Y. and Doi R.H. (2002) Cell-surface-anchoring role of N-terminal surface layer homology domains of *Clostridium cellulovorans* enge. J. Bacteriol. 184(4): 884–888.

Kraulis, J., Clore, G.M., Nilges, M., Jones, T.A., Pettersson, G., Knowles, J. and Gronenborn, A.M. (1989) Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei.* A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. Biochemistry *28*, 7241–7257.

Lamed, R., Setter, E. and Bayer, E.A. (1983) Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. J. Bacteriol. *156*, 828–836.

Lamed, R. and Bayer, E.A. (1988) The cellulosome of *Clostridium thermocellum*. Adv. Appl. Microbiol. *33*, 1–46.

Levitt, M. and Chothia, C. (1976) Structural patterns in globular proteins. Nature *261*, 552–557.

Li, H., Flora, R.M. and King, K.W. (1965) Individual roles of cellulase components derived from Trichoderma viride. Arch. Biochem. Biophys. *111*, 439–447.

Mandels, M. and Reese E.T. (1964) Fungal cellulases and the microbial decomposition of cellulosic fabric. Develop. Indust. Microbiol. *5*, 5–20.

Mandels, M., Weber, J. and Parizek, R. (1971) Enhanced cellulase production by a mutant of *Trichoderma viride*. Appl. Microbiol. *21*, 152.

McHale, A. and Coughlan, M.P. (1980) Synergistic hydrolysis of cellulose by components of the extracellular cellulase system of *Talaromyces emersonii*. FEBS Lett. *117*, 319.

Mechaly, A., Fierobe, H.P., Belaich, A., Belaich, J.P., Lamed, R., Shoham, Y. and Bayer, E.A. (2001) Cohesin-dockerin interaction in cellulosome assembly. A single hydroxyl group of a dockerin domain distinguishes between nonrecognition and high affinity recognition. J. Biol. Chem. 276(22):19678–19678.

Mitchinson, C. (2005) Genencor's (Genencor International, Inc., Palo Alto, CA) Perspective on the Key Barriers to Commercializing Biofuels at the 27th Symposium on Biotechnology for Fuels and Chemicals, Denver, May 2.

Montenecourt, B.S. and Eveleigh, D. E. (1979) Selective screening methods for the isolation of high yielding cellulase mutants of *Trichoderma reesei*. In: Hydrolysis of cellulose: Mechanism of Enzymatic and Acid Catalysis, Advances in Chemistry Series *181*, 289–301.

Morris, B. (1890) The germination of some of the gramineae. J. Chem. Soc. Lond. *57*, 497.

Newcombe, F.C. (1899) Cellulose-Enzymes. Annal. Bot. XIII, No. XLIX:49–81.

Notenboom, V., Boraston, A.B., Kilburn, D.G. and Rose, D.R. (2001) Crystal structures of the family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10A in native and ligand-bound forms. Biochemistry 40(21):6248–6256.

Orengo, C.A. and Thornton, J.M. (1993) Alpha plus beta folds revisited: some favoured motifs. Structure *1*, 105–120.

Orengo, C.A., Jones, D.T. and Thornton, J.M. (1994) Protein superfamilles and domain superfolds. Nature *372*, 631–634.

Petterson, L. G. (1975) The mechanism of enzymatic hydrolysis of cellulose. In Bailey, M., Enari, T.M. and Linko, M. (eds), Symposium on Enzymatic Hydrolysis of Cellulose. Helsinki, Finland, SITRA, pp. 255.

Pringsheim, H. (1912) Uber den fermentativen abbau der cellulose. Zeitschrift fuer physiologische Chemie *78*, 266–291.

Reese, E.T., Siu, R.G.H. and Levinson, H.S. (1950) The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. J. Bacteriol. *59*, 485.

Reinitzer, F. (1897) Ueber das zellwandlosende enzym der gerste. Hoppe-Seyler's Zeitschrift fur Physiologische Chemie *23*, 175–208.

Rincon, M. T., McCrae, S.I., Kirby, J., Scott, K.P. and Flint, H.J. (2001) Endb, A multidomain family 44 cellulase from *Ruminococcus flavefaciens* 17, binds to cellulose via a novel cellulose-binding module and to another R-flavefaciens protein via a dockerin domain. Appl. Environ. Microbiol. 67(10): 4426–4431.

Rouvinen, J., Bergfors, T., Teeri, T, Knowles, J. K. and Jones, T. A. (1990) Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. Science *249*, 380–386.

Sakon, J., Irwin, D., Wilson, D.B. and Karplus, A. (1997) Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. Nature Struct. Biol. 4(10):810–818.

Salamitou, S., Tokatlidis, K., Beguin, P. and Aubert J.P. (1992) Involvement of separate domains of the cellulosomal protein-S1 of *Clostridium thermocellum* in binding to cellulose and in anchoring of catalytic subunits to the cellulosome. Febs Letters 304(1):89–92.

Selby, K. (1969) The purification and properties of the C_1-component of the cellulase complex. In Hajny, G. J. and Reese, E. T. (eds), Cellulases and Their Applications. Advances in Chemistry Series *95*, 34–50.

Sigurskjold, B.W., Svensson, G.B. and Driguez, H (1994) Thermodynamics of ligand binding to the starch-binding domain of glucoamylase from *Aspergillus niger*. Eur. J. Biochem. 225(1):133–41.

Simpson, P.J., Xie, H.-F., Bolam, D.N., Gilbert, H.J. and Williamson, M.P. (2000) The structural basis for the ligand specificity of family 2 carbohydrate-binding modules. J. Biol. Chem. 275(52):41137–42.

Shoseyov, O., Takagi, M., Goldstein, M.A. and Doi, R.H. (1992) Primary sequence-analysis of *Clostridium cellulovorans* cellulose binding protein-A. Proc. Nat. Acad. Sci. USA 89(8):3483–3487.

Sorimachi, K., Jacks, A.J., Le Gal-Coeffect, M.-F., Williamson, G., Archer, D.B. and Willamson, M.P. (1996) Solution structure of the granular starch binding domain of glucoamylase from *Aspergillus niger* by nuclear magnetic resonance spectroscopy. J. Mol. Biol. 259(5):970–87.

Sorimachi, K., Le Gal-Coeffect, M. F., Williamson, G., Archer, D.B. and Williamson M.P. (1997) Solution structure of the granular starch binding domain of *Aspergillus niger* glucoamylase bound to beta-cyclodextrin. Structure 5(5):647–61.

Teeri, T.T. (1994) Hydrolysis of crystalline cellulose by native and engineered *Trichoderma reesei* cellulases. In Proceedings: The Symposium on Enzymic Degradation of Insoluble Polysaccharides, the 1994 Annual American Chemical Society Meeting, San Diego, CA.

Williamson, M.P., Le Gal-Coeffect, M.F., Sorimachi, K., Furniss, C.S.M., Archer, D.B. and Williamson, G. (1997) Function of conserved tryptophans in the *Aspergillus niger* glucoamylase 1 starch binding domain. Biochemistry 36(24):7535–9.

Wood, T.M. (1969) The cellulase of *Fusarium solani*, resolution of the enzyme complex. Biochem. J. *115*, 457.

Wood, T.M. (1988) Cellulase of *Trichoderma koningii*. Methods Enzymol. *160*, 221–234.

Wood, T.M. and McCrae, S.I. (1977) Cellulase from *Fusarium solani* purification and properties of the C-1 component. Carbohydr. Res. *57*, 117–133.

Wood, T.M. and McCrae, S.I. (1979) Synergism between enzymes involved in the solubilization of native cellulose. Brown, R.D. and Jurasek, L. (eds), Hydrolysis Of Cellulose: Mechanism Of Enzymatic And Acid Catalysis, Advances in Chemistry Series *181*, 181–209.

Woodward, J., Affholter, K.A., Kandy K., Noles, K.K., Troy, N.T. and Gaslightwala, S.F. (1992) Does cellobiohydrolase II core protein from *Trichoderma reesei* disperse cellulose macrofibrils? Enzyme and Microbial Technology *14*, 625–630.

Wooley R. and Ruth M. (1999) Conversion Of Hardwood Biomass To Ethanol Via Dilute Acid Cocurrent Prehydrolysis And Enzymatic SSCF – Process Design And Economics: Proceedings for the 21st Symposium on Biotechnology for Fuels and Chemicals, 6–04, Fort Collins, CO, May 2–6.

Perlack, R.D., Wright, L.L., Turhollow, A.F., Graham, R.L., Stokes, B.J. and Erbach, D.C. (2005) Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply", DOE Office of Energy Efficiency and Renewable Energy, Office of Biomass Program and the USDA Agriculture Research Service. *http://feedstockreview.ornl.gov/pdf/billion_ton_vision.pdf*

Xie, H.-F., Gilbert, H.J., Charnock, S.J., Davies, G.J., Willamson, M.P., Simpson, P.J., Raghotthama, S., Fontes, C.M. G.A., Dias, F.M.V., Ferreira, L.M.A. and Bolam, D.N. (2001) *Clostridium thermocellum* Xyn10B carbohydrate-binding module 22–2: the role of conserved amino acids in ligand binding. Biochemistry *40(31)*, 9167–9176.

Xu, Q., Bayer, E.A., Goldman, M., Kenig, R., Shoham, Y. and Lamed, R. (2004a) Architecture of the *Bacteroides cellulosolvens* cellulosome: description of a cell surface-anchoring scaffoldin and a family 48 cellulase. J. Bacteriol. *186*, 968–977.

Xu, Q., Barak, Y., Kenig, R., Shoham, Y., Bayer, E.A. and Lamed, R. (2004b) A novel *Acetivibrio cellulolyticus* anchoring scaffoldin that bears divergent cohesins. J. Bacteriol. *186*, 5782–5789.

CHAPTER 4

CELLULASES IN THE TEXTILE INDUSTRY

ARJA MIETTINEN-OINONEN*
VTT Biotechnology, Finland
**arja.miettinen-oinonen@stadia.fi*

1. INTRODUCTION

Cellulases are widely used in the textile industry for the manufacture and finishing of cellulose-containing materials. These enzymes are tools for improving basic processing steps in textile manufacture and creating new types of fabric. Their application in textile processing began in the 1980s with denim finishing, creating a fashionable stonewashed appearance in a process called biostoning (Kochavi *et al.*, 1990; Tyndall, 1990). In addition to biostoning, current commercial applications include biofinishing of cotton and other cellulose-based fibres and their use in detergents. In the detergent industry cellulases are used to provide cleaning and fabric-care benefits such as the brightening of colour in faded garments by removing fuzz (Maurer, 1997). The use of cellulases – and enzymes in general – in the textile industry confers a variety of advantages: enzymes are easy to use and treatments can be adapted to run on existing equipment and at different stages of textile wet processes; mild treatment conditions (*i.e.* temperature and pH) can be employed; enzymes are completely biodegradable and will not accumulate in the environment; enzymes are an economical option as they save chemicals and energy and can reduce processing times. Gene technology is widely used in the development of novel enzymes, the engineering of existing enzymes and for improvements in production efficiency. Apart from the conventional cellulase mixtures, cellulase products of tailored composition (*e.g.* enriched cellulase mixtures and monocomponent cellulases) are commercially available. Thus, by selecting different cellulase combinations a wide variety of effects on cellulose-containing materials can be achieved. In addition the performance of cellulases can be enhanced via product formulation by the incorporation of auxiliaries (*e.g.* surfactants) into the treatment liquor and by appropriate mechanical processing.

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 51–63

Cellulases account for approximately 14% of the world's industrial enzyme market and the current value of which is approximately 190 million US $ (Galante *et al*, 1998; Nierstrasz and Warmoeskerken, 2003). Approximately half of the enzymes marketed for textiles are cellulases (Ojapald, 2002; Nierstrasz and Warmoeskerken, 2003). In addition to the textile and detergent sectors, cellulases are also applied in the food, feed, and pulp and paper industries. A wide variety of bacteria and fungi produce cellulolytic enzymes of varying characteristics. Cellulases are hydrolytic enzymes and catalyse the breakdown of cellulose to smaller oligosaccharides and finally glucose. Cellulase activity refers to a multicomponent enzyme system consisting of three types of cellulases: (i) endoglucanases (EG: 1,4-β-D-glucan glucanohydrolase; EC 3.2.1.4), ii) cellobiohydrolases (also called exoglucanases, CBH: 1,4-β-D-glucan cellobiohydrolase; EC 3.2.1.91) and iii) β-glucosidases (BGL: cellobiase or β-D-glucoside glucohydrolase, EC 3.2.1.21). Endoglucanases cleave bonds along the length of the cellulose chains in the middle of the amorphous regions, resulting in a decrease in the degree of polymerisation (DP) of the substrate (reviewed in Teeri and Koivula, 1995; Teeri, 1997). Cellobiohydrolases are progressive enzymes, initiating their action from the ends of the cellulose chains. They attack the crystalline parts of the substrate, producing primarily cellobiose, and decrease the DP of the substrate very slowly. Cellobiohydrolases act synergistically with each other and with endoglucanases, thus mixtures of endoglucanases and cellobiohydrolases have greater activity than the sum of the activities of the individual enzymes acting alone. In the final cellulose hydrolysis step β-glucosidases hydrolyse the soluble oligosaccharides and cellobiose to glucose. Many of the fungal cellulases are modular proteins consisting of a catalytic domain, a carbohydrate-binding module (CBM) and a connecting linker. The role of the CBM is to mediate binding of the enzyme to the insoluble cellulose substrate.

Controlled hydrolysis by cellulases is used in textile processing to improve the surface properties and texture of cellulose-based fabrics. Advanced hydrolysis is not desired since this could cause too great a loss in fabric strength and weight. Using modern biotechnological tools different cellulase products having diverse cellulase profiles can be produced. Furthermore, novel techniques can improve the characteristics of enzymes, *e.g.* thermostability of cellulases (Voutilainen *et al*, 2004). By selecting a suitable cellulase product improved performance on different types of substrates can be achieved compared to that obtained with naturally occurring cellulases.

2. DENIM FINISHING

Denim is a cotton fabric woven with a dyed warp and raw white weft. Traditional blue denim jeans are dyed with indigo blue, and the stone-washed finish which gives a faded or worn appearance is achieved traditionally with pumice-stones. In large measure cellulases now replace the pumice stones to achieve a washed-out or aged appearance (Olson and Stanley, 1990). This process is called biostoning and is

currently the principle process used in the denim finishing industry. Approximately 1.8 thousand million pairs of jeans are produced annually and about 80% of these are finished using cellulases (Buchert and Heikinheimo, 1998). Denim washing efficiency is described as being 'high abrasion' due to the ability of cellulase to remove indigo from the material. In denim the indigo dye is attached to the surface of the yarn. In the biostoning process desized (removal of the starch coating) denim is treated with cellulases in a washing machine. The cellulases partially hydrolyse the surface of the fibre where the dye is bound. Since mechanical action is needed to remove the dye biostoning is usually carried out in jet or rotating drum washers. Typical treatment conditions are: temperature between 40–65°C, pH 4.5–7, treatment time 15–60 min, liquid ratio 1:3–1:15. The use of cellulases instead of stones has several advantages: (i) it prevents damage both to the washing machine and the garments; (ii) it eliminates the need for disposal of used stones; (iii) wastewater quality is improved; (iv) it eliminates the need for labour-intensive removal of dust from the finished garments, and (v) it permits increasing the garment load by 50% since no stones need to be added to the washing machine.

The cellulases used in denim finishing come from variety of sources (Table 1). Most are of fungal origin but bacterial and actinomycete cellulases have also been studied in relation to denim treatment (van Beckhoven *et al.*, 1996; Farrington *et al.*, 2001; van Solingen *et al.*, 2001). Cellulases for denim washing have traditionally been classified by the pH optimum of the enzyme: neutral cellulases operate in the pH range 6–8, and acid cellulases in the range of pH 4.5–6 (Videbaek *et al.*, 1994; Klahorst *et al.*, 1994; Auterinen *et al.*, 2004). Acid cellulases commercially used in biostoning mainly originate from the fungus *Trichoderma reesei.* One reason for the wide use of *T. reesei* cellulases is their low price. Acid cellulases also act aggressively on denim and result in abrasion over short washing times. Neutral cellulases by comparison are generally characterized by less aggressive action on cotton and the need for longer washing times (Klahorst *et al.*, 1994; Solovjeva *et al.*, 1998). The pH range of the currently used neutral cellulases is generally broader than that of the acid cellulases hence there is less need to control the pH of the treatment liquor when using neutral cellulases.

In addition to its source, the composition of a cellulase preparation affects denim-washing performance (Gusakov *et al.*, 1998, 2000; Heikinheimo *et al.*, 2000; Table 1). Whilst endoglucanases are needed for good abrasion, no direct correlation has been shown between abrasion level and any specific cellulase activity (Gusakov *et al.*, 2000). Several compositions have been proposed for obtaining good denim washing effects (Table 1). For example, of the principle cellulases of *T. reesei*, endoglucanase II has been shown to be the most effective at removing colour from denim (Heikinheimo *et al.*, 1998). By increasing the relative amount of endoglucanase II in a cellulase mixture processing times can be shortened resulting in more time- and cost-effective procedures (Miettinen-Oinonen and Suominen, 2002). Besides cost-effective treatments processes that preserve strength properties are essential in denim washing. Since cellulases hydrolyse cellulose the application of cellulases in denim wash or biofinishing (see below) often results in textile strength and weight losses. Much research has been directed to find out the

Table 1. Cellulases studied and used in denim finishing and their special performance

Source	Cellulase*	Application pH	Special performance	Reference
Trichoderma reesei	No or low CBHI	4.5–5.5	Low strength loss	Clarkson *et al.,* 1992a, b
	EG:CBH, 5:1		Low strength loss	Clarkson *et al.,* 1992c 1994
	Enriched CBHI		Low strength loss	Clarkson *et al.,* 1993
	EGII (purified)		Low hydrolysis level	Heikinheimo and Buchert, 2001
	EGIII + truncated EG and CBH		Decreased backstaining	Fowler *et al.,* 2001
Thielavia terrestris	EG	5	Almost bleached appearance, good abrasion	Schülein *et al.,* 1996, 1998
Chrysosporium lucknowense	Whole cellulase, EG	5	High abrasion, prevention of backstaining	Sinitsyn *et al.,* 2001
Penicillium occitanis		5.5		Belghith *et al.,* 2001
Melanocarpus albomyces	EG, EG:CBH	5–7	High abrasion, low backstaining	Miettinen-Oinonen *et al.,* 2004; Haakana *et al.,* 2004
***Streptomyces* sp.**	EG	5 - 10		van Solingen *et al.,* 2001
Myceliophthora thermophila	EG, EGI variants	6	Low strength loss, high abrasion, enhanced activity in alkaline pH	Schülein *et al.,* 1996, 1998; Osten and Schülein, 1999
Humicola insolens	EGI, V, EGI + V	6–7	Good abrasion, low strength loss, streak- reducing	Schülein *et al.,* 1998; Lund, 1997
***Acremonium* sp.**	EG	7	Low temperature, high abrasion	Schülein *et al.,* 1996, 1998
Fusarium oxysporium	EGI	n.r.	Low strength loss, little abrasion	Schülein *et al.,* 1998
Macrophomina phaseolina	EGV	n.r.	Good abrasion	Schülein *et al.,* 1998
Crinipellis scabela	EGV	n.r.	Good abrasion	Schülein *et al.,* 1998

n.r. = not reported

*EG = endoglucanase, CBH = cellobiohydrolase. The Roman numeral in front of EG or CBH refers to the individual endoglucanase or cellobiohydrolase.

choice of cellulase or cellulase mixtures and other process parameters that produce optimal results whilst retaining the strength of the fabric (Table I, Lenting and Warmoeskerken, 2001). A number of commercial cellulase products are currently available on the market each having its specific properties and yielding different results in denim washing.

During cellulase washing the released indigo dye tends to redeposit on the surface of the denim fabric resulting in colouring of the weft and re-colouring of the warp. This phenomenon is termed backstaining. Backstaining is an undesired property because the contrast between the blue and white yarn is reduced. Backstaining of the dye onto the pocket parts of a denim garment is a specific particular problem. Many studies have been undertaken to elucidate the mechanism of backstaining and prevent it. In early reports backstaining was claimed to be dependent on pH (Kochavi *et al*., 1990). Further experiments indicated that the nature of the enzyme used in washing has an impact on backstaining. In general, neutral cellulases tend to result in less backstaining whereas *T. reesei* cellulases (acidic) are associated with high backstaining (Klahorst *et al*., 1994). Indigo-cellulase affinities and enzyme adsorption to the white yarn of denim fabric have been suggested to cause backstaining (Cavaco-Paulo *et al*., 1998; Gusakov *et al*., 1998, 2000; Campos *et al.*, 2000). Inhibition of backstaining can be achieved by the following procedures: (i) the use of cellulases with less specific activity on indigo or denim; (ii) tailoring the composition of the cellulase preparation to achieve reduced backstaining with efficient abrasion; (iii) using cellulases which do not contain a CBM (cellulose-binding motif, formerly CBD for cellulose-binding domain) or where the CBMs have been removed; (iv) the addition of protease during rinsing or at the end of the cellulase washing step; (v) addition of anti-redeposition chemicals or mild bleaching agent during the enzyme washing or rinsing steps, and (vi) the presence of lipase during cellulase treatment (Tyndall, 1990; Cavaco-Paulo *et al*., 1998; Andreaus *et al.*, 2000; Yoon *et al*., 2000; Fowler *et al*., 2001; Uyama and Daimon, 2002; Miettinen-Oinonen *et al*., 2004; Haakana *et al*., 2004; Table I).

3. BIOFINISHING

Cellulases can also be exploited in fabric and garment finishing to produce higher value products. Cellulase treatment for finishing of cellulose-containing textile materials such as cotton, linen, hemp, lyocell, rayon and viscose materials is called biofinishing or biopolishing (Videbaek and Andersen, 1993). The most important parameters affecting successful biofinishing are the type of cellulases present in the enzyme preparation, the type of fibre being processed and the machinery used.

3.1. Cotton Finishing

In the biofinishing of cotton cellulases carry out a controlled surface hydrolysis. The fibre ends (microfibres) protruding from the fabric surface are weakened by cellulase action and are subsequently separated from the material with the aid of

Table 2. The benefits of biofinishing of yarn, fabric and garments

Performance	**Reference**
Cleared surface structure by reduced fuzz	Tyndall 1992; Pedersen *et al.,* 1992
Permanent decrease in pilling propensity	Pedersen *et al.,* 1992
Decreased hairiness	Pere *et al.,* 2001
Increased evenness of yarn	Pere *et al.,* 2001
Improved textile softness	Tyndall 1992; Pedersen *et al.,* 1992
Improved drapeability	Pedersen *et al.,* 1992; Kumar *et al.,* 1997
Brighter colours of the textile	Kumar *et al.,* 1997
Improved dimensional stability	Cavaco-Paulo 2001; Cortez *et al.,* 2002
Fashionable wash-down effects	Kumar *et al.,* 1997

mechanical action. The benefits of cellulase treatment of yarn, fabric and garment are listed in Table 2. In most cases the treatments are carried out on garments and fabrics. Treatment of yarn for pilling control may be advantageous in overcoming the dust problems often encountered with biofinishing of knitted fabrics. Biofinishing can be carried out after any textile wet processing step, that preferred being after bleaching of the fabric (Fig. 1). Partial removal of the dye occurs if cellulase treatment is done after dyeing, and the colour of the fabric can change (Nierstrasz and Warmoeskerken, 2003). If biofinishing is carried out before dyeing slightly deeper shades can sometimes be observed (Cavaco-Paulo and Gübitz, 2003). The combination of biofinishing and dyeing by adding a cellulase enzyme at the beginning of a dye cycle has also been reported (Ankeny, 2002). In this system cellulases acting at neutral pH are preferred and the performance of the enzyme in the dye bath depends on the dye.

The successfulness of cotton biofinishing is influenced by a number of parameters: pH, temperature, liquor ratio, enzyme concentration, time, mechanical agitation and machine type, fabric and fibre type, product quality, desired effect and cellulase composition (Cavaco-Paulo *et al.*, 1998; Liu *et al.*, 2000; Auterinen *et al.*, 2004). Improved performance is usually obtained when non-ionic surfactants and dispersing agents are present during the process (Traore and Buschle-Diller, 1999; Nierstrasz and Warmoeskerken, 2003); hard water, high ionic strength buffers and ionic surfactants have negative effects on cellulase performance (Cavaco-Paulo and Gübitz, 2003). Cellulases need to be inactivated after the treatment by raising the temperature and/or pH, washing the fabric with detergents or performing bleaching of the fabric in order to avoid undesirable strength and weight losses.

Figure 1. General stages of cotton wet processing

Commercial cellulases for biofinishing mainly originate from the fungi *T. reesei* and *Humicola insolens* (Lund and Pedersen, 1996; Galante *et al.*, 1998; Azevedo *et al.*, 2000; Cavaco-Paulo and Gübitz, 2003). Several studies have been conducted to evaluate the best cellulase component or cellulase mixture for high performance in biofinishing with minimal effects on the weight and strength properties of the fabric. In this regard, endoglucanases are the key enzymes in biofinishing. However, certain endoglucanases are known to negatively affect fabric strength. Results with individual *T. reesei* cellulases have shown that purified EGI and II caused greater loss of strength than purified CBHI but also had positive effects on the bending behaviour and pilling properties of cotton fabrics (Heikinheimo *et al.*, 1998). Furthermore EGII was good in pilling removal at low levels of hydrolysis and EGII-based cellulase mixtures gave positive depilling effects (Heikinheimo and Buchert, 2001; Miettinen-Oinonen *et al.*, 2001). Whole cellulase mixture was the best composition for cotton when considerable surface cleaning was required. However, endo-enriched cellulase resulted in reduced strength loss (Kumar *et al.*, 1997). Cellulase mixtures free of CBHI or CBHI-rich mixtures led to decreased strength loss compared to the whole mixtures (Clarkson *et al.*, 1992a-c, 1993). Strength loss can be minimized by using a monocomponent endoglucanase along with sufficient levels of mechanical action (Liu *et al.*, 2000; Lenting and Warmoeskerken, 2001). Furthermore, monocomponent endoglucanase has been shown to achieve high depilling with less weight loss compared to traditional whole acid cellulases (Liu *et al.*, 2000).

Sufficient mechanical agitation (shear force and mixing) is essential for successful biofinishing (Liu *et al.*, 2000; Cortez *et al.*, 2001; Cavaco-Paulo and Gübitz, 2003). Biofinishing has been introduced in industry in batch mode but not in continuous processes due to the lack of sufficient mechanical action (Aehle, 2004). Increasing mechanical agitation, *e.g.* using a jet-dyeing machine instead of a winch-dyer, has been shown to favour the attack of certain cellulase compositions (EG-rich cellulase products) compared to other types of composition (CBH-rich or whole mixtures), indicating that in addition to the nature of the cellulase composition biofinishing result is also dependent on the machine-type used (Cavaco-Paulo *et al.*, 1998; Cortez *et al.*, 2001).

3.2. Finishing of Man-made Cellulose Fibres

Biofinishing can also be used for processing man-made cellulose fibres such as viscose and the polynosic fibre lyocell (Kumar and Harnden, 1999; Ciechańska *et al.*, 2002; Carrillo *et al.*, 2003). Lyocell is a relatively new fibre invented in the early 1990s and is produced from wood pulp in a solvent spinning process (Courtaulds, 1995). Lyocell has high strength in both wet and dry states and is characterized by its tendency to fibrillate in the wet state as a result of abrasion. Cellulases have an essential role to play in removing this fibrillation. If the fibrils are not removed the surface of finished garments tends to exhibit high pilling and colour changes. The fibrillation of lyocell can also be used to engineer a variety

of surface finishes and optical effects such as "peach skin" and "mill-was" (Kumar and Harnden, 1998; Gandhi *et al*, 2002). To obtain the "peach skin" appearance cellulases are used to remove those fibrils formed during the primary fibrillation step which is performed at high temperature in alkaline solution. In the secondary fibrillation step after enzyme cleaning a peach skin appearance, in which the surface of the fabric consists of relatively short fibrils, is generated by washing or by dyeing. Conventionally the peach skin effect has been obtained using a three step batchwise process. Recently a novel method involving fibrillation, dyeing and enzyme cleaning in a single bath has been developed resulting in savings in treatment time (Gandhi *et al.*, 2002).

Cellulase products containing the whole range of cellulases and endo-enriched compositions have reported to be the optimal cellulases for defibrillation of lyocell (Aehle, 2004; Auterinen *et al*, 2004). Since lyocell is a strong fibre it retains its strength in cellulase treatments much better than other fibres (Auterinen, 2004). Mechanical action and its intensity also have a significant impact on the defibrillation of lyocell (Kumar and Harnden, 1998; Aehle, 2004).

4. OTHER APPLICATIONS

Apart from the well-established use of cellulases in the finishing of cellulose-based fibres their application in other areas of the textile industry such as in the preparatory processes of cotton and in the modification of bast fibres has also been studied. Cellulases have also been found to increase the alkaline solubility of treated pulp, and alkali soluble cellulose has been obtained using specific cellulase compositions (Vehviläinen *et al*, 1996; Rahkamo *et al*, 1996). The cellulose thus obtained can be utilized in developing new environmentally friendly processes for manufacturing cellulosic articles such as films, sponges and fibres.

4.1. Cotton Scouring

The purpose of cotton preparation (desizing, scouring and bleaching, Fig. 1) is to remove impurities, *e.g.* pectins, proteins and waxes, and prepare fabric for dyeing and any other wet processing treatments that follow. Scouring as a preparative step aims to produce absorbent fibre for uniform dyeing and finishing and is traditionally carried out by alkaline boiling. Pectinases, proteases, cellulases, xylanases and lipases have been studied for their potential application in enzymatic scouring and improved wettability has been obtained (reviewed in Aehle, 2004). Whilst enzymatic treatment of cotton with cellulases results in an absorbent fibre, weight and strength loss are incurred (Etters, 1999). Cellulase promotes the efficiency of cotton scouring with pectinase, lipase and protease but cannot function independently (Li and Hardin, 1998; Sangwatanaroj *et al*, 2003). Recently a commercial enzymatic scouring (bioscouring) treatment utilizing alkaline pectate lyase with a subsequent hot rinse in the presence of surfactants and chelators has been introduced to the market.

Seed-coat fragments derive from the outer layer of the cotton seed and need to be eliminated or bleached during the preparation of cotton. Seed coats are dark in colour and appear as dark spots in the fabric if still present during dyeing. Higher concentrations of chemicals are needed for the removal of seed coat fragments during scouring compared to other impurities. Cellulases have been shown to have potential for the removal of seed coat fragments during this process. Penetration of the alkaline solution and the degradation of seed coat fragments were increased after cellulase treatment (Csiszár *et al*., 1998). Additionally, cellulases were also found to degrade the small fibres attaching the seed coat fragments to fabrics thus reducing the amount of seed coat in the fabric. When treated with cellulases and other hydrolases seed coat fragments were hydrolysed faster than the cotton fabric suggesting that direct enzymatic removal of seed coat fragments might be possible (Csiszár *et al*., 2001).

4.2. Processing of Bast Fibres

Cellulases can also be used for the biofinishing of linen and other bast fibres. *Trichoderma* endoglucanases improve the pilling properties of linen fabric and the bending of flax fibres (Buschle-Diller *et al*., 1994; Pere *et al*., 2000). The chemical and structural properties of linen, such as the crystallinity of cellulose, are different from those in cotton. That the mode of action of cellulases is dependent on substrate, the effects obtained with linen can thus be different from those of cotton (Pere *et al.*, 2000). Greater weight and strength losses occur at lower cellulase dosages in linen treatments compared to cotton treatments. Thus the optimisation of cellulase treatments of linen, as regards cellulase composition, dosage and treatment time needs to be done with great care.

Retting of flax or other bast plants is a process where fibres are separated from the non-fibre tissues. Retting has been a major limitation for efficient flax fibre production. Water retting was the principal method but currently dew retting is that most utilised. The use of enzymes in retting has been studied for many years in order to obtain a more controlled way of isolating fibres and reducing effluents. Several enzyme products comprising mixtures of different enzymes such as pectinases, hemicellulases and cellulases have been tested in enzymatic retting (reviewed in Akin *et al*., 1997). Removal of pectin as the binder between cells is important in retting, hence pectinases have been the most effective enzymes in retting processes (Adamsen *et al*., 2002). The use of cellulases has been studied in up-grading of bast fibres for helping in further processing (Cavaco-Paulo and Gübitz, 2003). Good quality fibres have been obtained by enzymatic retting but so far this has not replaced commercial dew retting, one reason being the high cost (Akin *et al*., 2002).

ACKNOWLEDGEMENTS

I wish to thank Johanna Buchert from VTT Biotechnology for valuable comments on the manuscript and Mee-Young Yoon and Anna-Liisa Auterinen from Genencor Intl. for providing material for the preparation of this chapter.

REFERENCES

Adamsen, A., Akin, D. and Rigsby, L. (2002) Chelating agents and enzyme retting of flax. Textile Res. J. *72*, 296–302.

Aehle, W. (ed.) (2004) Enzymes in Industry. Production and Applications. Wiley-VCH, Weinheim, Germany.

Akin, D., Morrison, W., Gamble, G. and Rigsby, L. (1997) Effect of retting enzymes on the structure and composition of flax cell walls. Textile Res. J. *67*, 279–287.

Akin, D., Foulk, J. and Dodd, R. (2002) Influence on flax fibers of components in enzyme retting formulations. Textile Res. J. *72*, 510–514.

Andreaus, J., Campos, R., Gübitz, G. and Cavaco-Paulo, A. (2000) Influence of cellulase on indigo backstaining. Textile Res. J. *70*, 628–632.

Ankeny, M. (2002) Single-bath dyeing and bio-polishing. AATCC Review, May 2002, 16–19.

Auterinen, A-L., Jones, B. and Yoon, M-Y. (2004) Textile cellulases: searching molecules for improved performance. AATCC ICandE 2004.

Azevedo, H., Bishop, D. and Cavaco-Paulo, A. (2000) Effects of agitation level on the adsorption, desorption, and activities on cotton fabric of full length and core domains of EGV (*Humicola insolens*) and CenA (*Cellulomonas fimi*) Enzyme Microb. Technol. *27*, 325–329.

van Beckhoven, R., Lenting, H., Maurer, K-H. and Weiss, A. (1996) *Bacillus* cellulase and its applications. Pat. EP 0 739 982 A1.

Belghith, H., Ellouz-Chaabouni, and Gargouri, A. (2001) Biostoning of denims by *Penicillium occitanis* (Pol6) cellulases. J. Biotechnol. *89*, 257–262.

Buchert, J. and Heikinheimo, L. (1998) New cellulase processes for the textile industry. EU-project report. Carbohydr. Eur. *22*, 32–34.

Buschle-Diller, G., Zeronian, S., Pan, N. and Yoon, M-Y. (1994) Enzymatic hydrolysis of cotton, linen, ramie and viscose rayon fabrics. Textile Res. J. *64*, 270–279.

Campos, R., Cavaco-Paulo, A., Andreaus, J. and Gübitz, G. (2000) Indigo-cellulase interactions. Textile Res. J. *70*, 532–536.

Carrillo, F., Colom, X., López-Mesas, M., Lis, M., González, F. and Valldeperas, J. (2003) Cellulase processing of lyocell and viscose type fibres: kinetics parameters. Process Biochemistry *39*, 257–261.

Cavaco-Paulo, A., Almeida, L. and Bishop, D. (1998) Hydrolysis of cotton cellulose by engineered cellulases from *Trichoderma reesei*. Textile Res. J. *68*, 273–280.

Cavaco-Paulo, A. (2001) Improving dimensional stability of cotton fabrics with cellulase enzymes. Textile Res. J. *71*, 842–843.

Cavaco-Paulo, A. and Gübitz, G. (2003) Catalysis and processing. *In* Cavaco-Paulo, A. and Gübitz, G. (eds.) Textile processing with enzymes, Woodhead Publishing Ltd., England: 120–157.

Ciechańska, D., Struszczyk, H., Miettinen-Oinonen, A. and Strobin, G. (2002) Enzymatic treatment of viscose fibres based woven fabric. Fibres and Textiles in Eastern Europe. October/December 2002.

Clarkson, K., Larenas, E. and Weiss, G. (1992a) Methods for treating cotton-containing fabrics with cellulase low in cellobiohydrolases. Pat. WO 92/17574.

Clarkson, K., Larenas, E. and Weiss, G. (1992b) Methods for treating cotton-containing fabrics with cellulase free of exo-cellobiohydrolase, and the treated fabrics. Pat. WO 92/17572.

Clarkson, K., Larenas, E. and Weiss, G. (1992c) Methods for treating cotton-containing fabrics with cellulase. Pat. WO 92/06183.

Clarkson, K., Collier, K., Lad, P. and Weiss, G. (1993) Methods for treating cotton-containing fabrics with CBHI enriched cellulase. Pat. WO 93/22428.

Clarkson, K., Larenas, E. and Weiss, G. (1994) Method for reducing lint generation during treatment of cotton-containing and non-cotton-containing cellulosic fabric. Pat. WO 94/23113.

Cortez, J., Ellis, J. and Bishop, D. (2001) Cellulase finishing of woven, cotton fabrics in jet and winch machines. J. Biotechnol. *89*, 239–245.

Cortez, J., Ellis, J. and Bishop, D. (2002) Using cellulases to improve the dimensional stability of cellulosic fabric. Textile Res. J. *72*, 673–680.

Courtaulds Fibres Ltd. UK. (1995), Tencel. What it is, how it is made and what it can do.

Csiszár, E. Szakács, G. and Rusznák, I. (1998) Combining traditional cotton scouring with cellulase enzyme treatment. Textile Res. J. *3*, 163–167.

Csiszár, E., Losonczi, A., Szakács, G., Rusznák, I., Bezúr, L. and Reicher, J. (2001) Enzymes and chelating agent in cotton pre-treatment. J. Biotechnol. *89*, 271–279.

Etters, J.N. (1999) Cotton preparation with alkaline pectinase: and environmental advance. Textile Chemist and Colorist and American Dyestuff Reporter. November *1999*, 33–36.

Farrington, G., Anderson, P., Bergquist, P., Daniels, R., Moreland, D., Morgan H. and Williams, D. (2001) Compositions and methods for treating cellulose containing fabrics using truncated cellulase enzyme compositions. Pat. US 6,294,366.

Fowler, T., Clarkson, K., Ward, M., Collier, K. and Larenas, E. (2001) Method and compositions for treating cellulose containing fabrics using truncated cellulase enzyme compositions. Pat. US 6,268,196.

Galante, Y., De Conti, A. and Monteverdi, R. (1998) Application of *Trichoderma* enzymes in the textile industry. *In* Harman, G. and Kubicek, C. (eds.) *Trichoderma* and *Gliocladium*, Enzymes, biological control and commercial applications, Vol. 2: 311–326.

Gandhi, K., Burkinshaw, S., Taylor, J. and Collins, G. (2002) A novel route for obtaining a "peach skin effect" on lyocell and its blends. AATCC Review, April 2002:48–52.

Gusakov, A., Sinitsyn, A., Berlin, A., Popova, N., Markov, A., Okunev, O., Tikhomirov, D. and Emalfarb, M. (1998) Interaction between indigo and adsorbed protein as a major factor causing backstaining during cellulase treatment of cotton fabrics. Appl. Biochem.Biotechnol. *75*, 279–293.

Gusakov, A., Sinitsyn, A., Berlin, A., Markov, A. and Ankudimova, N. (2000) Surface hydrophobic amino acid residues in cellulase molecules as a structural factor responsible for their high denim-washing performance. Enzyme Microb. Technol. *27*, 664–671.

Haakana, H., Miettinen-Oinonen, A., Joutsjoki, V., Mäntylä, A., Suominen, P. and Vehmaanperä, J. (2004) Cloning of cellulase genes from *Melanocarpus albomyces* and their efficient expression in *Trichoderma reesei.* Enzyme Microb. Technol. *34*(2), 159–167.

Heikinheimo, L. and Buchert, J. (2001) Synergistic effects of *Trichoderma reesei* cellulases on the properties of knitted cotton fabric. Textile Res. J. 71, 672–677.

Heikinheimo, L., Cavaco-Paulo, A., Nousiainen, P., Siika-aho, M. and Buchert, J. (1998) Treatment of cotton fabrics with purified *Trichoderma reesei* cellulases. JSDC *114*, 216–220.

Heikinheimo, A., Buchert, J., Miettinen-Oinonen, A. and Suominen, P. (2000) Treating denim fabrics with *Trichoderma reesei* cellulases. Textile Res. J. 70(11): 969–973.

Klahorst, S., Kumar, A. and Mullins, M. (1994) Optimizing the use of cellulase enzymes. Textile Chemist and Colorist *26*, 2.

Kochavi, D., Videbaek, T. and Cadroni, D. (1990) Optimizing processing conditions in enzymatic stonewashing. American Dyestuff Reporter, Sept. 26–28.

Kumar, A. and Harnden, A. (1998) Performance characterization of cellulose enzymes in garment processing of 100% Tencel and its blends. AATCC ICandE 1998.

Kumar, A. and Harnden, A. (1999) Cellulase enzymes in wet processing of lyocell and its blends. Textile Chemist and Colorist and American Dyestuff Reporter *1*, 37–41.

Kumar, A., Yoon, M-Y. and Purtell, C. (1997) Optimizing the use of cellulase enzymes in finishing cellulosic fabrics. Textile Chemist and Colorist *29*, 37–42.

Lenting, H. and Warmoeskerken, M. (2001) Guidelines to come to minimized tensile strength loss upon cellulase application. J. Biotechnol. *89*, 227–232.

Li, Y. and Hardin, I. (1998) Enzymatic scouring of cotton – Surfactants, agitation and selection of enzymes. Textile Chemist and Colorist *30*, 23–29.

Liu, J., Otto, E., Lange, N., Husain, P., Condon, B. and Lund, H. (2000) Selecting cellulases for biopolishing based on enzyme selectivity and process conditions. Textile Chemist and Colorist and American Dyestuff Reporter *32*, 30–36.

Lund, H. (1997) A process for combined desizing and "stone-washing" of dyed denim. Pat. WO 97/18286.

Lund, H. and Pedersen, H. (1996) A method of obtaining a cellulosic textile fabric with reduced tendency to pilling formation. Pat. WO 96/17994.

Maurer, K-H. (1997) Development of new cellulases. *In*: van Ee, J., Misset, O. and Baas, E. (eds.) Enzymes in detergency. Marcel Dekker, Inc., New York, Pp. 175–202.

Miettinen-Oinonen, A., Heikinheimo, L., Buchert, J., Morgado, J., Almeida, L., Ojapalo, P. and Cavaco-Paulo, A. (2001) The role of *Trichoderma reesei* cellulases in cotton finishing. AATCC Review, January: 33–35.

Miettinen-Oinonen, A. and Suominen, P. (2002) Enhanced production of *Trichoderma reesei* endoglucanases and use of the new cellulase preparations in producing the stonewashed effect on denim fabric. Appl. and Env. Microbial. *8*, 3956–3964.

Miettinen-Oinonen, A., Londesborough, J., Joutsjoki, V., Lantto, R. and Vehmaanperä, J. (2004) Three cellulases from *Melanocarpus albomyces* for textile treatment at neutral pH. Enzyme Microb. Technol. 34/3-*4*, 332–341.

Nierstrasz, V.A. and Warmoeskerken, M. (2003) Process engineering and industrial enzyme applications. *In* Cavaco-Paulo, A. and Gübitz, G. (eds.) Textile processing with enzymes, Woodhead Publishing Ltd., England: 120–157.

Ojapalo, P. (2002) Tekstiilientsyymit – kaupalliset sovellukset. VTT Tekstiilitutkimusseminaari, Tampere, Finland, March 2002.

Olson, L. and Stanley, P. (1990) Compositions and methods to vary colour density. Pat. WO 90/02790.

Osten von der, C. and Schülein, M. (1999) Alkaline cellulases. Pat. US 5,912,157.

Pedersen, G., Screws, G. and Cedroni, D. (1992) Biopolishing of cellulosic fabrics. Can. Textile J., December: 31–35.

Pere, J., Miettinen-Oinonen, A., Heikinheimo, L., Puolakka, A., Nousiainen, P. and Buchert, J. (2000) Modification of cotton and linen fabrics with *Trichoderma reesei* cellulases. AATCC ICandE 2000: The New Millennium of textiles, Winston-Salem, NC. 4p.

Pere, J., Puolakka, A., Nousiainen, P. and Buchert, J. (2001) Action of purified *Trichoderma reesei* cellulases on cotton fibers and yarn. J. Biotechnol. *89*, 247–255.

Rahkamo, L., Siika-Aho, M., Vehviläinen, M., Dolk, M., Viikari, L., Nousiainen, P. and Buchert, J. (1996) Modification of hardwood dissolving pulp with purified *Trichoderma reesei* cellulases. Cellulose *3*, 153–163.

Sangwatanaroj, U., Choonukulpong, K. and Ueda, M. (2003) Cotton scouring with pectinase and lipase/protease/cellulase. AATCC Review, May 2003.

Sinitsyn, A., Gusakov, A., Grishutin, S., Sinitsyna, O. and Ankudimova, N. (2001) Application of microassays for investigation of cellulase abrasive activity and backstaining. J. Biotechnol. *89*, 233–238.

Schülein, M., Andersen, L., Lassen, S., Kauppinen, S. and Lange, L. (1996) Novel endoglucanases. Pat. WO 96/29397.

Schülein, M., Kauppinen, M., Lange, L., Lassen, S., Andersen, L., Klysner, S. and Nielsen, J. (1998) Characterization of fungal cellulases for fiber modification. ACS Symposium Series, 687 (Enzyme Applications in Fiber Processing): 66–74.

van Solingen, P., Meijer, D., van der Kleij, W., Barnett, C., Bolle, R., Power, S. and Jones, B. (2001) Cloning and expression of an endocellulase gene from a novel streptomycete isolated from an East African soda lake. Extremophiles *5*, 333–341.

Solovjeva, I., Ben-Bassat, A., Burlingame, R. and Chernoglazov, V. (1998) *Chrysosporium* cellulase and methods of use. Pat. WO 98/15633.

Teeri, T. (1997) Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. Tibtech. *15*, 160–167.

Teeri, T. and Koivula, A. (1995) Cellulose degradation by native and engineered fundal cellulases. Carbohydr. Eur. *12*, 28–33.

Traore, M. and Buschle-Diller,G. (1999) Influence of wetting agents and agitation on enzymatic hydrolysis of cotton. Textile Chemist and Colorist and American Dyestuff Reporter *1*, 51–56.

Tyndall, M. (1990) Upgrading garment washing techniques. American Dyestuff Reporter, May, 22–30.

Tyndall, R. M. (1992) Application of cellulase enzymes to cotton fabrics and garments. Textile Chemist and Colorist *24*, 23.

Uyama, N. and Daimon, K. (2002) Redeposition or backstain inhibition during stonewashing process. US 2002/0066144 A1

Vehviläinen M., Nousiainen P., Struszczyk H., Ciechanska D., Wawro D. and East G. (1996) Celsol -Biotransformation of cellulose for fibre spinning. *In* Kennedy J.F., Phillips G.O. and Williams P.A. (eds.) The Chemistry and Processing of Wood and Plant Fibrous Materials. Woodhead Publishing Limited, Cambridge: 197–204.

Videbaek, T. and Andersen, L. (1993) A process for defuzzing and depilling cellulosic fabrics. Pat. WO 93/20278.

Videbaek, T., Fich, M. and Screws, G. (1994) The jeans effect comes into being. International Textile Bulletin - Dyeing, Printing and Finishing *40*, 25–29.

Voutilainen, S., Boer, H., Linder, M., Vehmaanperä, J. and Koivula, A. (2004) Expression of Melanocarpus albomyces cellobiohydrolase Cel7B in Saccharomyces cerevisiae and directed evolution to improve its thermostability. Summer Course Glycosciences 2004. Wageningen. The Netherlands, 28 June – 1 July 2004.

Yoon, M-Y., McDonald, H., Chu, K. and Garratt, C. (2000) Protease, a new tool for denim washing. Textile Chemist and Colorist and American Dyestuff Reporter *32*, 25–29.

CHAPTER 5

XYLANASES: MOLECULAR PROPERTIES AND APPLICATIONS

F. I. JAVIER PASTOR[1]*, ÓSCAR GALLARDO[1], JULIA SANZ-APARICIO[2] AND PILAR DÍAZ[1]

[1]*Department of Microbiology, Faculty of Biology, University of Barcelona, Spain and*
[2]*Instituto de Química-Física Rocasolano, CSIC, Madrid, Spain*
**fpastor@ub.edu*

1. INTRODUCTION

The plant cell wall is a highly organized network of lignocellulose, made up of cellulose and cross-linked glycans embedded in a gel matrix of pectic substances and reinforced with structural proteins and aromatic compounds. Cellulose and hemicelluloses are the major components of cell wall polysaccharides, with hemicelluloses representing up to 20–35% of the total lignocellulosic biomass (de Vries and Visser, 2001). The major hemicellulose in cereals and hardwoods is xylan, while the main hemicellulose in softwoods is galactoglucomannan. Other less abundant hemicelluloses include glucomannan, xyloglucan, arabinogalactan and arabinan, the latter polymers often being found as side chains of pectins (de Vries and Visser, 2001). The degradation of hemicelluloses is mostly carried out by micro-organisms that can be found either free in nature or as a part of the digestive tract of higher animals. The hydrolytic enzymes produced by these micro-organisms are the key components for the degradation of plant biomass and carbon flow in nature (Shallom and Shoham, 2003).

Xylan is a major structural component of plant cell walls and, after cellulose, is the second most abundant renewable polysaccharide in nature (Collins *et al*, 2005). It is the main hemicellulose in hardwoods from angiosperms and is less abundant in softwoods from gymnosperms, accounting for approximately 15–30% and 7–12% of their total dry weights, respectively (Wong *et al*., 1988). In woody tissues xylan is located mainly in the secondary cell wall where, together with lignin, it forms an amorphous matrix that includes and embeds cellulose microfibrils.

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 65–82

Xylan interacts with lignin and cellulose via covalent and non-covalent linkages, these interactions being of importance for both protecting the cellulose microfibrils against biodegradation and maintaining the structural integrity of cell walls.

Xylan is a complex polysaccharide composed of a backbone of β-1,4-linked xylopyranosyl residues that, depending on the plant source, can be variably substituted by side chains of arabinosyl, glucuronosyl, methylglucuronosyl, acetyl, feruloyl and *p*-coumaroyl residues. Although homoxylans have been found in some plants and seaweeds, in the latter case also containing xylose β-1,3 linkages, xylans containing exclusively xylose residues are not widespread in nature (Beg *et al.*, 2001). Xylan from most plant sources occurs as a heteropolysaccharide and the terms glucuronoxylan and glucuronoarabinoxylan are commonly used to describe xylan from hardwoods and softwoods, respectively. These two types of xylan have 4-*O*-methyl α-D-glucuronic acid residues attached to C-2 of the xylose backbone units. Hardwoods have this substitution on approximately 10% of the xylose residues, while in softwoods around 20% of xylose residues are branched with glucuronic acid. Softwood xylan is also substituted with α-L-arabinofuranose on C-3 of approximately 13% of the xylose backbone residues (Coughlan and Hazlewood, 1993). The degree of polymerization is variable among xylans, being greater in hardwoods (150–200) than in softwoods (70–130) (Kulkarni *et al.*, 1999). While xylan from softwoods is not acetylated, xylan from hardwoods is highly acetylated, this substitution occurring on around 70% of the xylose units at C-2, C-3 or both (Coughlan and Hazlewood, 1993). The presence of acetyl groups makes xylan significantly more soluble in water (Biely, 1985). Xylan from grasses is usually referred to as arabinoxylan because of its large content in arabinosyl residues, which are linked to xylose at C-2 or C-3 or both. This xylan is acetylated and also has glucuronic acid present, albeit at a lower content compared to hardwoods. Feruloyl and coumaroyl residues ester-linked to C-5 of arabinose side chains are found in xylans from different sources, and may be involved in the covalent cross-linking of xylan molecules with lignin or with other xylan molecules. As a consequence of all these features, xylans constitute a very heterogeneous group of polysaccharides showing microheterogeneity with respect to the degree and nature of branching in each category. Xylans containing rhamnose and galactose residues have also been described from different plant sources (Wong *et al.*, 1988).

2. ENZYMATIC DEGRADATION OF XYLAN

Due its heterogeneity and complex nature, the complete breakdown of xylan requires the action of a large variety of hydrolytic enzymes (Biely, 1985; Coughlan and Hazlewood, 1993). These enzymes can be classified into two main groups: those acting on the xylose backbone, and those cleaving the side chains. Degradation of the xylose backbone depends on xylanases, that cleave bonds within the polymer, and β-xylosidases that release xylose units from xylobiose and xylooligomers. Removal of xylan side chains is catalysed by α-L-arabinofuranosidases, α-D-glucuronidases, acetyl xylan esterases, ferulic acid esterases and *p*-coumaric acid esterases (Fig. 1). Xylan degradation is quite widespread among saprophytic

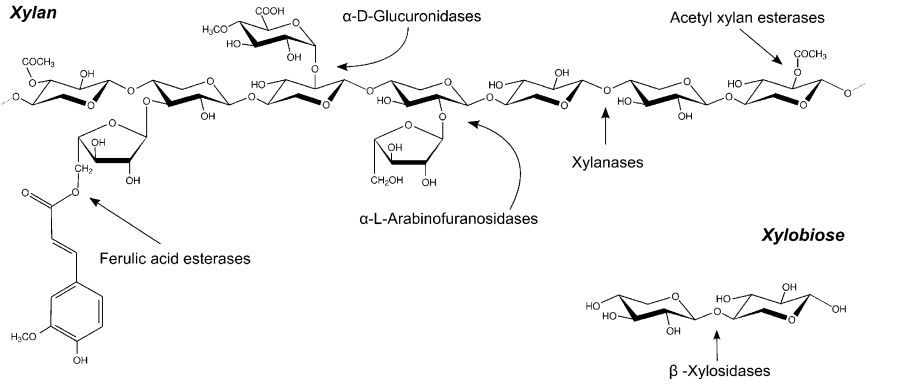


Figure 1. **Structure of xylan and the sites of attack by xylanolytic enzymes**. The backbone of xylan chains is composed of β-1,4-linked xylopyranose residues. This backbone can be variously substituted by side chains of arabinosyl, glucuronosyl, methylglucuronosyl, acetyl, feruloyl and *p*-coumaroyl residues. Hydrolysis of the xylan backbone is carried out by xylanases that hydrolyse internal linkages in xylan, and β-xylosidases that release xylose units from xylobiose and xylooligomers, while removal of xylan side chains is catalysed by α-L-arabinofuranosidases, α-D-glucuronidases, acetyl xylan esterases, ferulic acid esterases and *p*-coumaric acid esterases

micro-organisms, including bacteria and fungi, as well as in the rumen microbiota, that possess complete xylanolytic enzyme systems (Biely, 1985; Sunna and Antranikian, 1997; Krause *et al*, 2003). Synergism between xylan degrading enzymes has been extensively studied and found to frequently occur between xylanases and side chain cleaving enzymes, between xylanases and β-xylosidases, and also between different xylanases (Coughlan *et al*, 1993; de Vries *et al*, 2000). In this way, xylan degradation can proceed despite that the access of xylanases to their targets in the xylan backbone may be obstructed by side chain substituents and that these substituents may be more readily released from xylan fragments than from the polymeric substrate.

3. XYLANASES

3.1. Function, Expression and Multiplicity

Xylanases (endo-β-1,4-xylanases, EC 3.2.1.8) cleave the xylan backbone into smaller oligosaccharides. They are the key enzymes for xylan degradation and differ in their specificities toward the xylan polymer. Many cleave only at unsubstituted regions whereas others have a requirement for side chains in the vicinity of the cleaved bonds (Coughlan and Hazlewood, 1993). Most xylanases are also active on xylooligomers of degree of polymerization greater than 2, showing increasing affinity for xylooligomers of increasing length. Xylanases are endo type enzymes that hydrolyse internal linkages in xylan and act by a random attack mechanism yielding a mixture of xylooligosaccharides from the polymer.

Nevertheless, characterization of *Aeromonas* xylanases which produce only one oligosaccharide type as a reaction product from xylan suggested an alternative mode of hydrolysis: an exo type mechanism from one end of the polymer (Kubata *et al*., 1995) similar to the processive mode of action of exocellulases in cellulose degradation (Lynd *et al*., 2002). However, as this type of mechanism would first require depletion of the side chains from xylan, the occurrence of true exoxylanases for xylan degradation seems unlikely.

As xylan is a large polymer that cannot penetrate into cells, xylanases have to be secreted to the extracellular environment to reach and hydrolyse it. Generally, xylanases are induced in most micro-organisms during their growth on substrates containing xylan. Small soluble oligosaccharides released from xylan by the action of low levels of constitutively produced enzymes are transported inside cells where they induce xylanase expression (Kulkarni *et al*., 1999). In the fungus *Cryptococcus albidus*, xylobiose is considered to be the natural inducer or the direct precursor of compounds that induce xylanase expression (Biely, 1985). Regulation of xylanase synthesis is often coordinated with the expression of cellulases, as in the case of *Aspergillus* in which several regulatory proteins including the transcriptional activator XlnR and the carbon catabolite repressor CreA have been identified and characterized (de Vries and Visser, 2001). Although most xylanases are extracellular enzymes, usually secreted by the Sec-dependent pathway, periplasmic xylanases have been described in some rumen bacteria and in *Cellvibrio mixtus* (Fontes *et al*., 2000). These periplasmic xylanases are probably involved in the breakdown of large xylooligosaccharides and protected in this location from extracellular proteases. Recently, a cytoplasmic xylanase that may represent a new type of enzymes involved in xylan degradation has been characterized from *Paenibacillus barcinonensis* (Gallardo *et al*., 2003). In common with three other xylanases characterized from *Bacillus* and *Aeromonas*, the *P. barcinonensis* enzyme lacks a signal peptide for export outside the cytoplasm. These four enzymes constitute a group of highly homologous xylanases whose proposed role is the hydrolysis of the short oligosaccharides resulting from extracellular xylan hydrolysis once they have been transported inside cells.

Many xylan degrading micro-organisms produce a multiplicity of xylanases with different but overlapping specificities (Wong and Saddler, 1988). This has been evidenced for important xylanase producers like *Aspergillus*, *Trichoderma*, *Streptomyces* and *Bacillus* amongst others (Beg *et al*., 2001; Sunna and Antranikian, 1997). Multiplicity of xylanases can arise from different post-translational processing of the same gene product (Ruiz-Arribas *et al*., 1995), though very often several xylanase-encoding genes have been isolated from a defined microbial strain. In this regard, at least 4 xylanase genes have been isolated from *Fibrobacter succinogenes* (Jun *et al*., 2003) and 6 have been characterized in *Cellvibrio japonicus* (formerly *Pseudomonas cellulosa*) (Emami *et al*., 2002). The production of a multienzyme system of xylanases, in which each enzyme has a specific function, represents a strategy to achieve efficient hydrolysis of xylan.

Cellulosomes are secreted multienzyme complexes found in anaerobic bacteria that mediate the attachment between cells and cellulose particles to enhance the efficiency of cellulose and hemicellulose degradation (Bayer *et al.*, 2004). They contain numerous enzymes, most of them cellulases, kept together through interactions between pairs of dockerin and cohesin domains located on the enzymes and on a non-catalytic scaffolding protein. Xylanases are also found among the component enzymes of most known cellulosomes. Xylan degradation can be essential in making cellulose available for enzymatic hydrolysis, since cellulose is present in close proximity with xylan in the plant cell wall matrix. By analogy, the term xylanosome has been proposed for extracellular protein aggregates predominantly composed of xylanases reported in several bacteria, though these have not been as well characterized as cellulosomes (Beg *et al.*, 2001; Jiang *et al.*, 2005).

3.2. Molecular Architecture and Classification

According to their molecular architecture, xylanases, like cellulases and other carbohydratases, can be classified into two types: single domain and multidomain enzymes. Xylanases of the first type contain a single catalytic domain, whereas multidomaim xylanases have a modular structure that comprises, in addition to a catalytic domain, several ancillary domains joined by linker sequences (Gilkes *et al.*, 1991; Gilbert and Hazlewood, 1993). These domains may fold and function independently and can mediate binding to cellulose (cellulose binding domains, CBD), xylan (xylan binding domains), or to cellulosome scaffolding proteins (dockerin domains); or they may have functions that are not yet fully identified such as Fn3 domains (Kataeva *et al.*, 2002) or SLH domains (Ali *et al.*, 2001).

Cellulose binding domains, the most abundant non-catalytic domains in xylanases, have been characterized and grouped into several families (Tomme *et al.*, 1998). They promote binding to different forms of crystalline or amorphous cellulose and may disrupt cellulose microfibrils to facilitate degradation by cellulases (Linder and Teeri, 1997). Modules that mediate binding to xylans have also been characterized (Black *et al.*, 1995). They include previously designated thermostabilizing domains, found in some thermophilic (Lee *et al.*, 1993) and also in mesophilic xylanases (Blanco *et al.*, 1999), that have subsequently been shown to promote binding to xylan (Sunna *et al.*, 2000). Identification of domains that promote binding to other carbohydrates such as chitin and starch has prompted the term carbohydrate binding module to group all domains that mediate binding to carbohydrates (Boraston *et al.*, 2004).

The heterogeneity and complex nature of xylan has resulted in a diversity of xylanases with varying specificities, sequences and folds. Wong *et al.* (1988) classified xylanases into two types according to their physicochemical properties: a group of low molecular weight (<30 kDa) and basic pI, and a second group of high molecular weight (>30 kDa) and acidic pI. Although many xylanases fit in these groups, the number of xylanases described since then has increased enormously, and at present many xylanases of intermediate properties have been identified that

do not fit either of these categories. A more complete classification system was introduced in 1991 for the glycoside hydrolases (EC 3.2.1.x), a group of enzymes in which xylanases are included, which is based on primary structure comparison of catalytic domains and the grouping of enzymes in families of related sequences (Henrissat, 1991). From the 35 families identified in 1991, the classification has been updated regularly with newly sequenced enzymes and comprises over 100 families (GH1 to GH106) at present (Coutinho and Henrissat, 1999; Carbohydrate Active enZYmes CAZY server at *http://afmb.cnrs-mrs.fr/CAZY/*). Most families comprise enzymes with the same substrate specificity, though several families are polyspecific and include enzymes active on different carbohydrates. The finding of related structures in different families has resulted in the introduction of a higher hierarchical level of classification known as the superfamily or clan. A clan is a group of families that are believed to share a common ancestor and show related tertiary structures together with conservation of the catalytic residues and catalytic mechanism (Henrissat and Bairoch, 1996).

Xylanases are usually classified into glycoside hydrolase families 10 (formerly F) and 11 (formerly G). These two families include, respectively, xylanases of the high MW/low pI and low MW/high pI groups previously mentioned, but each of the families includes many other xylanases with physicochemical properties widely different from these two groups (Sunna and Antranikian, 1997; Beg *et al.*, 2001). Family 10 includes cellobiohydrolases (exocellulases) and endo-β-1,3-xylanases besides xylanases (endo-β-1,4-xylanases), while family 11 is monospecific, comprising solely xylanases. A small number of recently characterized xylanases do not show sequence similarity to families 10 or 11. Instead, these xylanases exhibit homology to enzymes belonging to glycoside hydrolase families 5, 7, 8 and 43. Accordingly, the group of families containing xylanases should be expanded to include these new enzymes (Collins *et al.*, 2005).

3.3. Structure

The three dimensional structures of many bacterial and fungal xylanases from families 10 and 11 have been reported. In addition, crystal structures from new xylanases belonging to families 5 and 8 have been solved within the last few years.

Glycoside hydrolase families 5 and 10 are members of clan GH-A which includes 17 glycoside hydrolase families. Despite large differences in size and sequence, members of this clan possess a catalytic domain of 250–450 amino acids which shares a common $(\alpha/\beta)_8$ TIM-barrel fold and a remarkable conservation of the 3D structure of the active site (Fig. 2a). Many family 10 enzymes are modular, containing a carbohydrate binding module connected to the catalytic domain by a flexible linker. To date, the only crystal structures of full length xylanases known for this family are those of Xyn10A from *Streptomyces olivaceoviridis* (Fujimoto *et al.*, 2000) which displays a small substrate binding domain linked by a Gly/Pro-rich region, and Xyn10C from *Cellvibrio japonicus* which bears a family 15 carbohydrate binding module (Pell *et al.*, 2004a). In both cases the linker is not visible

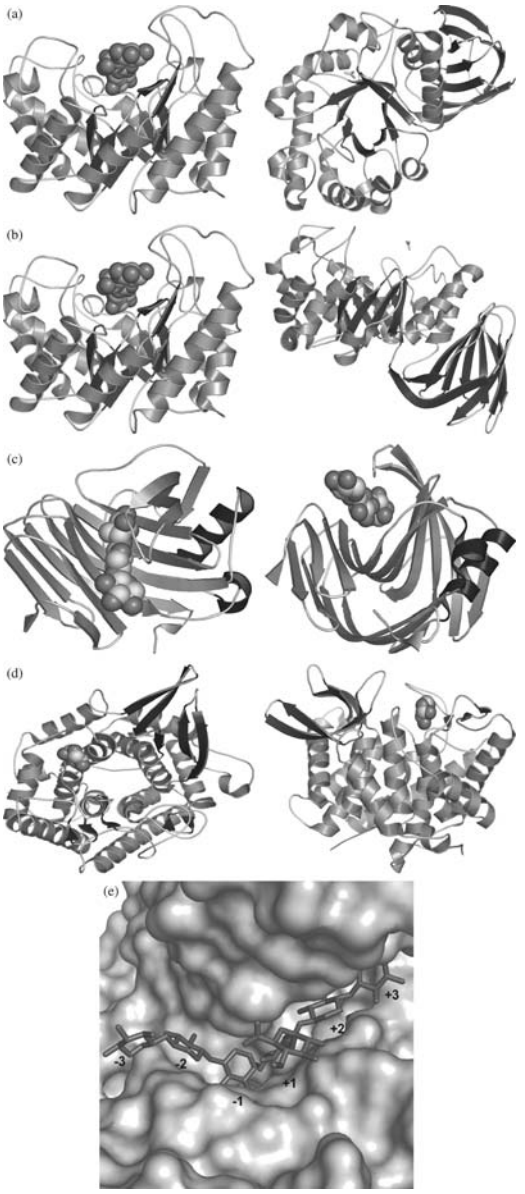


Figure 2. **Three-dimensional structure of xylanases and their complexes with xylooligosaccharides**. **a)** Structure of the catalytic domain of GH10 *Cellvibrio japonicus* (formerly *Pseudomonas fluorescens*) xylanase A in complex with xylopentaose showing the typical $(\alpha/\beta)_8$-barrel fold. The oligosaccharide chain is occupying subsites -1 to $+4$ within the active site cleft (Lo Leggio *et al*, 2000). **b)** Structure of GH5 xylanase A from *Erwinia chrysanthemi* showing the $(\alpha/\beta)_8$-barrel catalytic domain and a small β9-barrel domain, probably a xylan binding module (Larson *et al*, 2003). **c)** Structure of the E94A mutant of the GH11 xylanase from *Bacillus agaradhaerens* in complex with xylotriose. The typical

in the electron density maps. Attempts to crystallize other family 10 xylanases in their intact multidomain forms have been unsuccessful so far, possibly due to the mobility of these enzymes allowed by the flexibility of the linkers that connect the domains.

The three dimensional structure of the family 5 xylanase XynA from *Erwinia chrysanthemi* has been reported (Larson *et al.*, 2003). This enzyme contains a short module of 100 residues located at the C-terminus, that is similar to carbohydrate binding modules of family 20 and attributed to promote xylan binding (Fig. 2b). Comparison of XynA to the known catalytic domains of families 5 and 10 shows that XynA is no more structurally equivalent to family 5 than it is to family 10 xylanases (Larson *et al.*, 2003).

Family 11 xylanases usually have catalytic domains of 180–200 residues that fold into a β-sheet motif known as the β jelly-roll fold (Fig. 2c) and shared with family 12 cellulases, both members of clan GH-C. Interestingly, enzymes of family 11 are more specific for xylan and they usually do not contain additional domains, though some examples of this family such as TfxA from *Thermobifida fusca* show domains for substrate binding (Irwin *et al.*, 1994).

Finally, the structures of two family 8 enzymes, a xylanase from *Pseudoalteromonas haloplanktis* (Van Petegem *et al.*, 2003) and BH2105 enzyme from *Bacillus halodurans* that hydrolyses xylooligosaccharides but is not active on xylan (Honda and Kitaoka, 2004), have been reported. They fold into an $(\alpha/\alpha)_6$-barrel, common among other inverting glycosidases: family 9 endoglucanases, family 15 glucoamylases and family 48 cellobiohydrolases (Fig. 2d).

The active site of xylanases is an extended open cleft consistent with their endo mode of action. It usually displays between four and seven subsites for binding the xylopyranose rings in the vicinity of the catalytic site. The binding sites are numbered in either direction from the catalytic site and are assigned positive numbers in the direction of the reducing end of the substrate, which constitutes the leaving group (the aglycone), and negative numbers in the direction of the non-reducing end, which remains bound to the catalytic site in the intermediate complex state (Fig. 2e). The crystal structures of family 10 and family 11 xylanases in complex with oligosaccharides and inhibitors have revealed detailed information on how the xylan backbone binds to these enzymes (Lo Leggio *et al.*, 2000; Sabini *et al.*, 2001). The data show that residues forming subsites −2 and −1 are conserved in each family, while the aglycone moiety can be located in variable sites.

Figure 2. (Continued) fold of the family is a β-sheet motif known as the β jelly-roll. The xylose chain spans subsites −3 to −1 (Sabini *et al.*, 2001). **d)** Structure of the GH8 xylanase from *Pseudomonas haloplanktis* in complex with xylose, that occupies putative subsite +4. The polypeptide chain of this family folds into an $(\alpha/\alpha)_6$-barrel common to other inverting glycoside hydrolases (Van Petegem *et al.*, 2003). **e)** Molecular surface of the inactive xylanase 10B E262S mutant from *Cellvibrio mixtus* showing a close-up view of the active site tunnel. A xylotriose moiety is occupying subsites −3 to −1, while xylotriose at subsites +1 to +3 is decorated with 4-*O*-methyl glucuronic acid (Pell *et al.*, 2004b)

Comparisons of the catalytic properties of the xylanases in the two major families, 10 and 11, show that family 10 xylanases exhibit greater catalytic versatility or lower substrate specificity than enzymes in family 11, and can also exhibit activity on some cellulosic substrates such as aryl cellobiosides (Biely *et al*., 1997). Furthermore, family 10 xylanases show greater activity on short xylooligosaccharides than family 11 enzymes, indicating the existence of smaller binding sites. In agreement with this, enzymes from family 10 yield smaller hydrolysis products from glucuronoxylan and rhodymenan (β-1,3-β-1,4-xylan), further hydrolysing the oligosaccharides which are released from these polysaccharides by family 11 xylanases, and usually cleave xylans to a greater extent (Kolenová *et al*., 2005).

Substituents in the xylan chain seem to affect xylanases differently and appear to constitute a more serious steric hindrance for family 11 members. Indeed, one of the differences between the two major families of xylanases is that family 11 enzymes hydrolyse unsubstituted regions of xylan, whereas the corresponding family 10 xylanases are able to attack decorated regions of the polysaccharide. Recent studies on Xyn10B from *Cellvibrio mixtus* in complex with decorated xylooligosaccharides revealed that the two major decorations of xylan, arabinose and 4-*O*-methylglucuronic acid, can be accommodated in selected glycone and aglycone subsites of family 10 enzymes (Pell *et al*., 2004b) (Fig. 2e). A more recent crystal structure of a family 10 xylanase from *Thermoascus aurantiacus* in complex with an arabinofuranosyl-ferulate substrate has shown extensive interaction of the arabinose with the enzyme, thus suggesting a role for the xylan side chains as determinants of specificity for this family of xylanases (Vardakou *et al*., 2005).

3.4. Catalytic Mechanism

Glycoside hydrolases act by two major mechanisms which result in a net retention or inversion of the anomeric configuration (Rye and Withers, 2000; Collins *et al.*, 2005). Xylanases from families 10 and 11 catalyse hydrolysis by a double displacement mechanism with retention of the anomeric configuration. Two conserved glutamate residues suitably located in the active site (approximately 5.5 Å apart) are the catalytically active residues. One of the glutamate residues acts as a general acid catalyst that protonates the glycosidic oxygen, while the second performs a nucleophilic attack resulting in the departure of the leaving group and the formation of an α-glycosyl/enzyme intermediate. In a second step, the first catalytic residue now functions as a general base, abstracting a proton from a water molecule that attacks the anomeric carbon and hydrolyses the glycosyl/enzyme intermediate. This second substitution at the anomeric carbon generates a product with the same stereochemistry as the substrate, thus retaining the anomeric configuration (Collins *et al.*, 2005). A similar catalytic mechanism is found in family 5 enzymes. By contrast, family 8 glycoside hydrolases operate with inversion of the anomeric configuration. Glutamate and aspartate are believed to be the catalytic residues. One acts as a general acid catalyst while the other acts as a general base catalyst, and are typically separated in the active centre by a distance of around 9.5 Å to allow the

accommodation of a water molecule between the anomeric carbon and the general base. As a consequence of the single displacement mechanism, the configuration of the anomeric centre is inverted.

4. β-XYLOSIDASES AND DEBRANCHING ENZYMES

4.1. β-xylosidases

As mentioned above, efficient degradation of xylans requires not only the action of xylanases but also the cooperation of other enzymes such as β-xylosidases and chain degrading enzymes. β-xylosidases (β-1,4-xylosidases, EC 3.2.1.37) hydrolyse xylobiose and short chain xylooligosaccharides generated by the action of xylanases, releasing xylose from the non-reducing end. The affinity of β-xylosidases for oligosaccharides decreases with the increasing degree of polymerisation of the latter. These enzymes do not usually hydrolyse xylan but they can hydrolyse artificial substrates such as *p*-nitrophenyl-β-D-xylopyranoside which is frequently used as a substrate for routine colourimetric assays of β-xylosidase activity (Coughlan and Hazlewood, 1993). β-xylosidases are grouped into glycoside hydrolase families 3, 39, 43 and 52, including enzymes with inverting and retaining catalytic mechanisms (Shallom and Shoham, 2003). Many β-xylosidases also show transxylosidase activity, allowing the formation of products of higher molecular weight than the starting substrates and hence the production of novel xylose-containing substances under appropriate conditions. This suggests a possible application of these enzymes in the synthesis of specific oligosaccharides (de Vries and Visser, 2001). As regards the location of β-xylosidases, they appear to be mainly cell-associated, though many extracellular β-xylosidases have also been reported (Coughlan *et al*., 1993).

4.2. α-L-arabinofuranosidases

α-L-arabinofuranosidases (EC 3.2.1.55) are exo-acting enzymes that catalyse the cleavage of terminal arabinose residues from the side chains of xylan and other arabinose-containing polysaccharides (Saha, 2000). They have been classified into families 43, 51, 54 and 62 of the glycoside hydrolases and are usually assayed colourimetrically by monitoring the hydrolysis of *p*-nitrophenyl-α-L-arabinofuranoside (Coughlan *et al*., 1993). The apparent release of arabinose by some xylanases gave rise to a classification of xylanases into debranching and non-debranching enzymes depending on whether or not they produced free arabinose in addition to cleaving the xylan backbone (Matte and Forsberg, 1992). However, it seems that the reported release of arabinose by xylanases may have been due to contamination of these enzymes by trace amounts of arabinofuranosidases. Indeed, synergistic activity between xylanases and arabinofuranosidases makes it possible that a small amount of contaminant may yield detectable amounts of free arabinose (Coughlan *et al*., 1993).

4.3. α-D-glucuronidases

These enzymes (EC 3.2.1.131) hydrolyse the linkages between 4-*O*- methylglucuronic/glucuronic acid and xylose residues in glucuronoxylan, and are found exclusively in glycoside hydrolase family 67. Despite their role in the biodegradation of xylan, there are not many examples of these enzymes. Some show activity only on short xylooligomers or small model molecules, while others can release glucuronic acid from polymeric xylan (Puls, 1992).

4.4. Acetyl Xylan Esterases

Acetyl xylan esterases (EC 3.1.1.72) remove the acetyl groups from acetylated xylan. These enzymes are a late discovery due to the lack of appropriate substrates. Although xylans can be highly acetylated, most of the substrates used to study enzymatic degradation were obtained by alkali extraction, a method that tends to strip the acetyl groups from xylan (Sunna and Antranikian, 1997). Acetyl xylan esterases play an important role in the hydrolysis of xylan since acetyl groups can hinder the approach of enzymes that cleave the xylan backbone, hence the removal of these substituents facilitates the action of xylanases (Coughlan *et al.*, 1993).

4.5. Hydroxycinnamic Acid Esterases

Ferulic acid and *p*-coumaric acid esterases cleave the ester bonds between arabinose side chains and feruloyl or *p*-coumaroyl residues, respectively (Williamson *et al.*, 1998). These residues can cross-link xylan molecules to each other or to lignin. The high yield of ferulic and *p*-coumaric esterases produced by the fungus *Neocallimastix* and other rumen anaerobic fungi seems to provide these microorganisms with an advantage over bacteria by conferring on them the ability to degrade and utilise phenolic ester-linked arabinoxylans (Borneman *et al.*, 1993).

5. APPLICATIONS OF XYLANASES

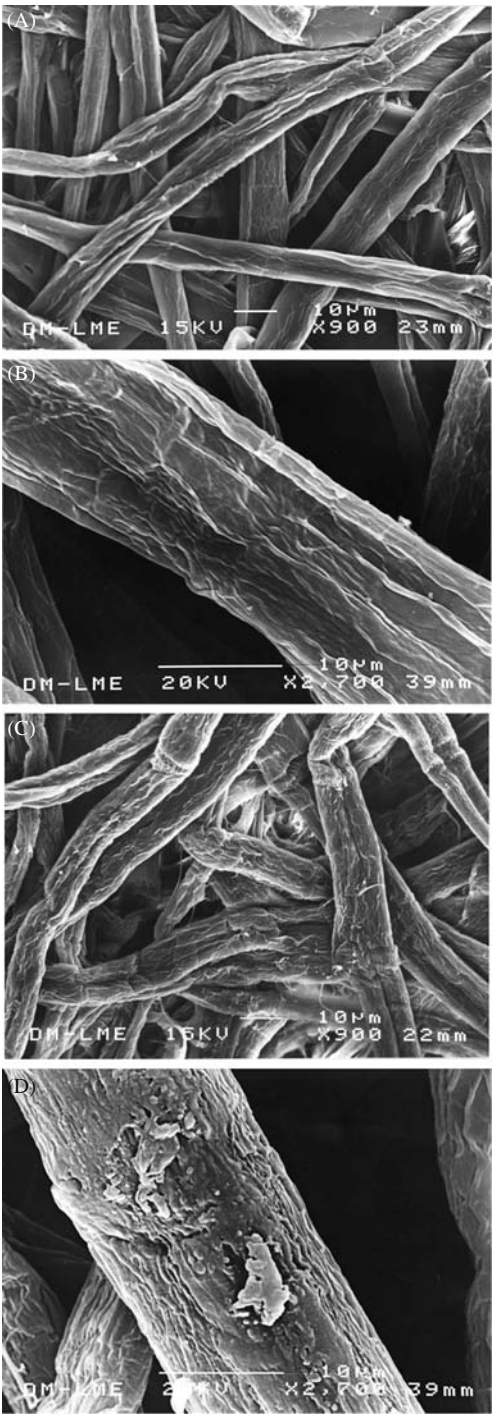
Microbial hemicellulases, especially xylanases, have important applications in industry due to their enormous potential to modify and transform the lignocellulose and cell wall materials abundant in vegetal biomass which is used in a wide variety of industrial processes. The biotechnological application of xylanases began in the 1980s in the preparation of animal feed, and later expanded to the food, textile and paper industries. Since then the biotechnological use of these enzymes has increased dramatically, covering a wide range of industrial sectors. At present, xylanases together with cellulases and pectinases account for 20% of the global industrial enzyme market (Polizeli *et al.*, 2005).

Xylan is present in large amounts in wastes from the agricultural and food industries. Xylanases are thus of increasing importance for the bioconversion of lignocellulosic biomass, including urban solid residues, to xylose and other fermentable sugars for the production of biological fuels (ethanol) (Lee, 1997). Bioconversion of xylan to the low calorie sweetener xylitol is a promising field where xylanases can also play a key role (Polizeli *et al*., 2005). Other less well documented potential applications of xylanases include their use as additives in detergents, in the preparation of plant protoplasts, the production of pharmacologically active oligosaccharides as antioxidants, and the use of xylanases possessing transxylosidase activity for the synthesis of new surfactants (Bhat, 2000; Collins *et al*., 2005).

Xylanases are used as additives in animal feeds for monogastric animals, together with cellulases, pectinases and many other depolymerizing enzymes. Enzyme degradation of arabinoxylans, commonly found as ingredients of feeds, reduces the viscosity of the raw materials thus facilitating better mobility and absorption of other components of the feed and improving nutritional value (Polizeli *et al*., 2005). The incorporation of xylanase into the rye- or wheat-based diets of broiler chickens resulted in an improvement in weight gain of chicks and their feed conversion efficiency (Bedford and Classen, 1992). Similar improvements can be obtained for pigs fed on a wheat-based diet supplemented with xylanases and phospholipases (Diebold *et al*., 2005).

The application of xylanases along with pectinases in the juice and wine industries facilitates the extraction and clarification of the final products (Bhat, 2000). These enzymes can also increase the stability of fruit pulp and release aroma precursors. As regards the latter, a recombinant yeast strain expressing a fungal xylanase produced a wine with increased fruity aroma (Ganga *et al*., 1999). Xylanases can be also used in brewing to reduce beer's haze and viscosity, and to increase wort filterability (Polizeli *et al*., 2005). As baking additives, xylanases degrade flour hemicelluloses resulting in a redistribution of water from pentosans to gluten, thus giving rise to an increase in bread volume and crumb quality, and an antistaling effect (Linko *et al.*, 1997). This can be further enhanced when amylases are used in combination with xylanases (Monfort *et al*., 1996).

The major current industrial application of xylanases is in the pulp and paper industry where xylanase pretreatment facilitates chemical bleaching of pulps, resulting in important economic and environmental advantages over the non-enzymatic process (Viikari *et al*., 1994; Bajpai, 2004). Xylanases do not remove lignin-based chromophores directly but instead degrade the xylan network that traps the residual lignin. Degradation of xylan in xylan-lignin complexes or reprecipitated on the surface of fibres after kraft cooking, allows a more efficient extraction of lignin by the bleaching chemicals. Microscopic analysis of pulps shows that xylanase treatment opens up fibre surface which exhibits detached material, in contrast to the smooth surface of untreated fibres (Fig. 3) (Roncero *et al*., 2000). Xylanase-boosted bleaching results in up to 20–25% savings on chlorine-based chemicals and a reduction of 15–20% in the generation of pollutant organic chlorine compounds from lignin degradation (adsorbable organic halogens, AOX)


(A)
DM-LME 15KV 10μm X900 23mm
(B)
DM-LME 20KV 10μm X2,700 39mm
(C)
DM-LME 15KV 10μm X900 22mm
(D)
DM-LME 10μm X2,700 39mm

(Viikari *et al.*, 1994; Bajpai, 2004). The reduction in the amount of chemical bleaching agents required to obtain a target paper brightness has contributed to the replacement of elemental chlorine by the less polluting chlorine dioxide in elemental chlorine free (ECF) bleaching sequences, or to the total replacement of chlorine compounds by alternative bleaching agents such as hydrogen peroxide and ozone in total chlorine free (TCF) bleaching sequences.

The bleaching efficiency of different fungal and bacterial xylanases has been analysed. Although many of the enzymes tested are highly efficient as bleaching aids, notable differences can appear depending on the family and traits of each particular enzyme (Elegir *et al.*, 1995; Clarke *et al.*, 1997). The response to enzyme-aided bleaching can also be affected by the bleaching sequence, wood species concerned and the pulping method (Suurnäkki *et al.*, 1996; Nelson *et al.*, 1995; Christov *et al.*, 2000). At present, many microbial xylanases are available on the market and are successfully used in pulp mills (Beg *et al.*, 2001).

In relation to the bleaching process, xylanase treatment can modify pulp-refining properties. In some cases, enzymatically treated pulps require greater beating, while the strength properties of the paper are not affected or only slightly modified (Roncero *et al.*, 2003; Vicuña *et al.*, 1995). A decrease in xylan content by enzyme treatment has been reported to modify the ageing and brightness reversion of pulps and paper, which can show increased stability and less yellowing tendency after enzyme treatment (Buchert *et al.*, 1997).

Besides xylanases, other hemicellulases have also been tested as bleaching aids with various results. Among them, β-mannanases have been shown to facilitate bleaching, eliminating residual lignin and increasing paper brightness, though the effect of mannanases is usually less pronounced than that of xylanases (Montiel *et al.*, 1999; Bhat, 2000). Advances in understanding lignin degradation has resulted in the proposal of a different strategy for bleaching, involving the direct removal of lignin by lignin depolymerizing enzymes (laccases and peroxidases). Laccases from several fungi and from *Streptomyces* have been successfully assayed (Bourbonnais *et al.*, 1997; Sigoillot *et al.*, 2005; Arias *et al.*, 2003) whereas few examples of brightness improvement with manganese peroxidases have been reported to date.

The application of xylanases in the pulp and paper industry is not restricted to bleaching. The good results obtained in this field have stimulated the evaluation of the use of xylanases in other stages of pulp and paper manufacture. Application of xylanases in mechanical pulping, pulp drainage or the deinking of recycled fibres is currently being evaluated, and the promising results obtained are leading to an expanding use of xylanases in this industry and an increasing importance for xylanases in the world enzyme market.

Figure 3. **SEM analysis of cellulose fibres**. Scanning electron micrographs of fully unbleached **(A)** and **(C)**, or oxygen delignified **(B)** and **(D)** *Eucalyptus* kraft pulps before or after xylanase treatment. **(A)** and **(B)** untreated pulps showing fibres with smooth surfaces; **(C)** and **(D)** xylanase treated pulps showing flakes and filaments of material detached from the fibre surface. Courtesy of Dr. T. Vidal (Roncero *et al.*, 2000)

REFERENCES

Ali, M.K., Kimura, T., Sakka, K. and Ohmiya, K. (2001) The multidomain xylanase Xyn10B as a cellulose-binding protein in *Clostridium stercorarium*. FEMS Microbiol. Lett. *198*, 79–83.

Arias, M.E., Arenas, M., Rodríguez, J., Soliveri, J., Ball, A.S. and Hernández, M. (2003) Kraft pulp biobleaching and mediated oxidation of a nonphenolic substrate by laccase from *Streptomyces cyaneus* CECT 3335. Appl. Environ. Microbiol. *69*, 1953–1958.

Bajpai, P. (2004) Biological bleaching of chemical pulps. Crit. Rev. Biotechnol. *24*, 1–58.

Bayer, E.A., Belaich, J.P., Shoham, Y. and Lamed, R. (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annu. Rev. Microbiol. *58*, 521–554.

Bedford, M.R. and Classen, H.L. (1992) The influence of dietary xylanase on intestinal viscosity and molecular weight distribution of carbohydrates in rye-fed broiler chicks. In: Visser, J., Beldman, G., Kusters-van Someren, M.A., Voragen, A.G.J. (eds) Xylans and xylanases. Elsevier, Amsterdam, pp. 361–370.

Beg, Q.K., Kapoor, M., Mahajan, L. and Hoondal, G.S. (2001) Microbial xylanases and their industrial applications: a review. Appl. Microbiol. Biotechnol. *56*, 326–338.

Bhat, M.K. (2000) Cellulases and related enzymes in biotechnology. Biotechnol. Adv. *18*, 355–383.

Biely, P. (1985) Microbial xylanolytic systems. Trends Biotechnol. *3*, 286–290.

Biely, P., Vršanská, M., Tenkanen, M. and Kluepfel, D. (1997) Endo-β-1,4-xylanase families: differences in catalytic properties. J. Biotechnol. *57*, 151–166.

Black, G.W., Hazlewood, G.P., Millward-Sadler, S.J., Laurie, J.I. and Gilbert, H.J. (1995) A modular xylanase containing a novel non-catalytic xylan-specific binding domain. Biochem. J. *307*, 191–195.

Blanco, A., Díaz, P., Zueco, J., Parascandola, P. and Pastor, F.I.J. (1999) A multidomain xylanase from a *Bacillus* sp. with a region homologous to thermostabilizing domains of thermophilic enzymes. Microbiology *145*, 2163–2170.

Boraston, A.B., Bolam, D.N., Gilbert, H.J. and Davies, G.J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. *382*, 769–781.

Borneman, W.S., Ljungdahl, L.G., Hartley, R.D. and Akin, D.E. (1993) Feruloyl and *p*-coumaroyl esterases from the anaerobic fungus *Neocallimastix* strain MC-2: properties and functions in plant cell wall degradation. In: Coughlan M.P and Hazlewood G.P (eds) Hemicelluloses and hemicellulases. Portland Press, London and Chapel Hill, pp. 85–102.

Bourbonnais, R., Paice, M.G., Freiermuth, B., Bodie, E. and Borneman, S. (1997) Reactivities of various mediators and laccases with kraft pulp and lignin model compounds. Appl. Environ. Microbiol. *63*, 4627–4632.

Buchert, J., Bergnor, E., Lindblad, G., Viikari, L. and Ek, M. (1997) Significance of xylan and glucomannan in the brightness reversion of kraft pulps. Tappi J. 80(6):165–171.

Clarke, J.H., Rixon, J.E., Ciruela, A., Gilbert, H.J. and Hazlewood, G.P. (1997) Family-10 and family-11 xylanases differ in their capacity to enhance the bleachability of hardwood and softwood paper pulps. Appl. Microbiol. Biotechnol. *48*, 177–183.

Collins, T., Gerday, C. and Feller, G. (2005) Xylanases, xylanase families and extremophilic xylanases. FEMS Microbiol. Rev. *29*, 3–23.

Coughlan, M.P. and Hazlewood, G.P. (1993) β-1,4-D-Xylan-degrading enzyme systems: biochemistry, molecular biology and applications. Biotechnol. Appl. Biochem. *17*, 259–289.

Coughlan, M.P., Tuohy, M.G., Filho, E.X.F., Puls, J., Claeyssens, M., Vrsanská, M. and Hughes, M.M. (1993) Enzymological aspects of microbial hemicellulases with emphasis on fungal systems. In: Coughlan M.P and Hazlewood G.P (eds) Hemicelluloses and hemicellulases. Portland Press, London and Chapel Hill, pp. 53–84.

Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. In: Gilbert, H.J., Davies, G.S., Henrissat, B., Svensson. B. (eds) Recent Advances in Carbohydrate Bioengineering. The Royal Society of Chemistry, Cambridge, pp. 3–12.

Christov, L., Biely, P., Kalogeris, E., Christakopoulos, P., Prior, B.A. and Bhat, M.K. (2000) Effects of purified endo-β-1,4-xylanases of family 10 and 11 and acetyl xylan esterases on eucalypt sulfite dissolving pulp. J. Biotechnol. *83*, 231–244.

Diebold, G., Mosenthin, R., Sauer, W.C., Dugan, M.E.R. and Lien, K.A. (2005) Supplementation of xylanase and phospholipase to wheat-based diets for weaner pigs. J. Anim. Physiol. Anim. Nutr. *89*, 316–325.

De Vries, R.P., Kester, H.C.M., Poulsen, C.H., Benen, J.A.E. and Visser, J. (2000) Synergy between enzymes from *Aspergillus* involved in the degradation of plant cell wall polysaccharides. Carbohydr. Res. *327*, 401–410.

De Vries, R.P. and Visser, J. (2001) *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microbiol. Mol. Biol. Rev. *65*, 497–522.

Elegir, G., Sykes, M. and Jeffries, T.W. (1995) Differential and synergistic action of *Streptomyces* endoxylanases in prebleaching of kraft pulps. Enzyme Microb. Technol. *17*, 954–959.

Emami, K., Nagy, T., Fontes, C.M.G.A., Ferreira, L.M.A. and Gilbert, H.J. (2002) Evidence for temporal regulation of the two *Pseudomonas cellulosa* xylanases belonging to glycoside hydrolase family 11. J. Bacteriol. *184*, 4124–4133.

Fontes, C.M.G.A., Gilbert, H.J., Hazlewood, G.P., Clarke, J.H., Prates, J.A.M., McKie, V.A., Nagy, T., Fernandes, T.H. and Ferreira, L.M.A. (2000) A novel *Cellvibrio mixtus* family 10 xylanase that is both intracellular and expressed under non-inducing conditions. Microbiology *146*, 1959–1967.

Fujimoto, Z., Kuno, A., Kaneko, S., Yoshida, S., Kobayashi, H., Kusakabe, I. and Mizuno, H. (2000) Crystal structure of *Streptomyces olivaceoviridis* E-86 beta-xylanase containing xylan-binding domain. J. Mol. Biol. *300*, 575–585.

Gallardo, O., Diaz, P. and Pastor, F.I.J. (2003) Characterization of a *Paenibacillus* cell-associated xylanase with high activity on aryl-xylosides: a new subclass of family 10 xylanases. Appl. Microbiol. Biotechnol. *61*, 226–233.

Ganga, M.A., Piñaga, F., Vallés, S., Ramón, D. and Querol, A. (1999) Aroma improving in microvinification processes by the use of a recombinant wine yeast strain expressing the *Aspergillus nidulans* xlnA gene. Int. J. Food Microbiol. *47*, 171–178.

Gilbert, H.J. and Hazlewood, G.P. (1993) Bacterial cellulases and xylanases. J. Gen. Microbiol. *139*, 187–194.

Gilkes, N.R., Henrissat, B., Kilburn, D.G., Miller Jr., R.C. and Warren, R.A.J. (1991) Domains in microbial ß-1,4-glycanases: sequence conservation, function, and enzyme families. Microbiol. Rev. *55*, 303–315.

Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. *280*, 309–316.

Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. Biochem. J. *316*, 695–696.

Honda, Y. and Kitaoka, M. (2004) A family 8 glycoside hydrolase from *Bacillus halodurans* C-125 (BH2105) is a reducing end xylose-releasing exo-oligoxylanase. J. Biol. Chem. 279, 55097:55103.

Irwin, D., Jung, E.D. and Wilson, D.B. (1994) Characterization and sequence of a *Thermomonospora fusca* xylanase. Appl. Environ. Microbiol. *60*, 763–770.

Jiang, Z.Q., Deng, W., Li, X.T., Ai, Z.L., Li, L.T. and Kusakabe, I. (2005) Characterization of a novel, ultra-large xylanolytic complex (xylanosome) from *Streptomyces olivaceoviridis* E-86. Enzyme Microb. Technol. *36*, 923–929.

Jun, H.S., Ha, J.K., Malburg Jr., L.M., Gibbins, A.M.V. and Forsberg, C.W. (2003) Characteristics of a cluster of xylanase genes in *Fibrobacter succinogenes* S85. Can. J. Microbiol. *49*, 171–180.

Kataeva, I.A., Seidel III, R.D., Shah, A., West, L.T., Li, X.L. and Ljungdahl, L.G. (2002) The fibronectin type 3-like repeat from the *Clostridium thermocellum* cellobiohydrolase CbhA promotes hydrolysis of cellulose by modifying its surface. Appl. Environ. Microbiol. *68*, 4292–4300.

Krause, D. O., Denman, S.E., Mackie, R.I., Morrison, M., Rae, A.L., Attwood, G.T. and McSweeney, C.S. (2003) Opportunities to improve fiber degradation in the rumen: microbiology, ecology, and genomics. FEMS Microbiol. Rev. *27*, 663–693.

Kolenová, K., Vršanská. M. and Biely, P. (2005) Purification and characterization of two minor endo-β-1,4-xylanases of *Schizophyllum commune*. Enzyme. Microb. Technol. *36*, 903–910.

Kubata, B.K., Takamizawa, K., Kawai, K., Suzuki, T. and Horitsu, H. (1995) Xylanase IV, an exoxylanase of *Aeromonas caviae* ME-1 which produces xylotetraose as the only low-molecular-weight oligosaccharide from xylan. Appl. Environ. Microbiol. *61*, 1666–1668.

Kulkarni, N., Shendye, A. and Rao, M. (1999) Molecular and biotechnological aspects of xylanases. FEMS Microbiol. Rev. *23*, 411–456.

Larson, S.B., Day, J., Barba de la Rosa, A.P., Keen, N.T. and McPherson, A. (2003) First crystallographic structure of a xylanase from glycoside hydrolase family 5: Implications for catalysis. Biochemistry 42, 8411:8422.

Lee, J. (1997) Biological conversion of lignocellulosic biomass to ethanol. J. Biotechnol. *56*, 1–24.

Lee, Y.E., Lowe, S.E., Henrissat, B. and Zeikus, J.G. (1993) Characterization of the active site and thermostability regions of endoxylanase from *Thermoanaerobacterium saccharolyticum* B6A-RI. J. Bacteriol. *175*, 5890–5898.

Linder, M. and Teeri, T.T. (1997) The roles and function of cellulose-binding domains. J. Biotechnol. *57*, 15–28.

Linko, Y.Y., Javanainen, P. and Linko, S. (1997) Biotechnology of bread baking. Trends Food Sci. Technol. *8*, 339–344.

Lo Leggio, L., Jenkins, J., Harris, G.W., Pickersgill, R.W. (2000) X-ray crystallographic study of xylopentaose binding to *Pseudomonas fluorescens* xylanase A. Proteins *41*, 362–373.

Lynd, L.R., Weimer, P.J., van Zyl, W.H. and Pretorius, I.S. (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. *66*, 506–577.

Matte, A. and Forsberg, C.W. (1992) Purification, characterization, and mode of action of endoxylanases 1 and 2 from *Fibrobacter succinogenes* S85. Appl. Environ. Microbiol. *58*, 157–168.

Monfort, A., Blasco, A., Prieto, J.A. and Sanz, P. (1996) Combined expression of *Aspergillus nidulans* endoxylanase X24 and *Aspergillus oryzae* α-amylase in industrial bakers's yeasts and their use in bread making. Appl. Environ. Microbiol. *62*, 3712–3715.

Montiel, M.D., Rodríguez, J., Pérez-Leblic, M.I., Hernández, M., Arias, M.E. and Copa-Patiño, J.L. (1999) Screeening of mannanase in actinomycetes and their potential application in the biobleaching of pine kraft pulps. Appl. Microbiol. Biotechnol. *52*, 240–245.

Nelson, S.L., Wong, K.K.Y., Saddler, J.N. and Beatson, R.P. (1995) The use of xylanase for peroxide bleaching of kraft pulps derived from different softwood species. Pulp Paper Can. 96 (7):42–45.

Pell, G., Szabo, L., Charnock, S.J., Xie, H., Gloster, T.M., Davies, G.J. and Gilbert, H.J. (2004a) Structural and biochemical analysis of *Cellvibrio japonicus* xylanase 10C. J. Biol. Chem. *279*, 11777–11788.

Pell, G., Taylor, E.J., Gloster, T.M., Turkenburg, J.P., Fontes, C.M.G.A., Ferreira, L. M. A., Nagy, T., Clark, S. J., Davies, G.J. and Gilbert, H. J. (2004b) The mechanisms by which family 10 glycoside hydrolases bind decorated substrates. J. Biol. Chem. *279*, 9597–9605.

Polizeli, M.L.T.M., Rizzatti, A.C.S., Monti, R., Terenzi, H.F., Jorge, J.A. and Amorim, D.S. (2005) Xylanases from fungi: properties and industrial applications. Appl. Microbiol. Biotechnol. *67*, 577–591.

Puls, J. (1992) α-glucuronidases in the hydrolysis of wood xylans. In: Visser, J., Beldman, G., Kusters-van Someren, M.A., Voragen, A.G.J. (eds) Xylans and xylanases. Elsevier, Amsterdam, pp. 213–224.

Roncero, M.B., Torres, A.L., Colom, J.F. and Vidal, T. (2000) Effects of xylanase treatment on fibre morphology in totally chlorine free bleaching (TCF) of *Eucalyptus* pulp. Process Biochem. *36*, 45–50.

Roncero, M.B., Torres, A.L., Colom, J.F. and Vidal, T. (2003) Effect of xylanase on ozone bleaching kinetics and properties of *Eucalyptus* kraft pulp. J. Chem. Technol. Biotechnol. *78*, 1023–1031

Ruiz-Arribas, A., Fernández-Abalos, J.M., Sánchez, P., Garda, A.L. and Santamaría, R.I. (1995) Overproduction, purification, and biochemical characterization of a xylanase (Xys1) from *Streptomyces halstedii* JM8. Appl. Environ. Microbiol. *61*, 2414–2419.

Rye, C.S. and Whithers, S.G. (2000) Glycosidase mechanisms. Curr. Opin. Chem. Biol. *4*, 573–580.

Sabini, E., Wilson, K.S., Danielsen, S., Schulein, M. and Davies, G.J. (2001) Oligosaccharide binding to family 11 xylanases: both covalent intermediate and mutant product complexes display $(^{2,5})$B conformations at the active centre. Acta Crystallogr. D57:1344–1347.

Saha, B.C. (2000) α-L-arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. Biotechnol. Adv. *18*, 403–423.

Shallom, D. and Shoham, Y (2003) Microbial hemicellulases. Curr. Op. Microbiol. *6*, 219–228.

Sigoillot, C., Camarero, S., Vidal, T., Record, E., Asther, M., Pérez-Boada, M., Martínez, M.J., Sigoillot, J.C., Asther, M., Colom, J.F. and Martínez, A.T. (20005) Comparison of different fungal enzymes for bleaching high-quality paper pulps. J. Biotechnol. *115*, 333–343.

Sunna, A. and Antranikian, G. (1997) Xylanolytic enzymes from fungi and bacteria. Crit. Rev. Biotechnol. *17*, 39–67.

Sunna, A., Gibbs, M.D. and Bergquist, P.L. (2000) The thermostabilizing domain, XynA, of *Caldibacillus cellulovorans* xylanase is a xylan binding domain. Biochem. J. *346*, 583–586.

Suurnäkki, A., Clark, T.A., Allison, R.W., Viikari, L. and Buchert, J. (1996) Xylanase- and mannanase-aided ECF and TCF bleaching. Tappi. J. 79(7):111–117.

Tomme, P., Boraston, A., McLean, B., Kormos, J., Creagh, A.L., Sturch, K., Gilkes, N.R., Haynes, C.A., Warren, R.A.J. and Kilburn, D.G. (1998) Characterization and affinity applications of cellulose-binding domains. J. Chromatogr. B *715*, 283–296.

Van Petegem, F., Collins, T., Meuwis, M.A., Gerday, C., Feller, G. and van Beeumen, J. (2003) The structure of a cold-adapted family 8 xylanase at 1.3 Å resolution. J. Biol. Chem. *278*, 7531–7539.

Vardakou, M., Flint, J., Christakopoulos, P., Lewis, R.J., Gilbert, H.J. and Murray, J.W. (2005) A family 10 *Thermoascus aurantiacus* xylanase utilizes arabinose decorations of xylan as significant substrate specificity determinants. J. Mol. Biol. *352*, 1060–1067.

Vicuña, R., Oyarzún, E. and Osses, M. (1995) Assessment of various commercial enzymes in the bleaching of radiata pine kraft pulps. J. Biotechnol. *40*, 163–168.

Viikari, L., Kantelinen, A., Sundquist, J. and Linko, M. (1994) Xylanases in bleaching: from an idea to the industry. FEMS Microbiol. Rev. *13*, 335–350.

Williamson, G., Kroon, P.A. and Faulds, C.B. (1998) Hairy plant polysaccharides: a close shave with microbial esterases. Microbiology. *144*, 2011–2023.

Wong, K.K.Y., Tan, L.U.L. and Saddler, J.N. (1988) Multiplicity of β-1,4-xylanase in micro-organisms: functions and applications. Microbiol. Rev. *52*, 305–317.

CHAPTER 6

MICROBIAL XYLANOLYTIC CARBOHYDRATE ESTERASES

EVANGELOS TOPAKAS AND PAUL CHRISTAKOPOULOS*

Biotechnology Laboratory, Chemical Engineering Department, National Technical University of Athens, Greece

**hristako@chemeng.ntua.gr*

1. INTRODUCTION

The plant cell wall represents the most abundant reservoir of organic carbon in the biosphere with 10^{11} tons synthesized annually. The degradation of this macromolecular structure by microbial enzymes is a key biological process that is central to the carbon cycle, herbivore nutrition and host invasion by phytopathogenic fungi and bacteria. The plant cell wall also represents an important industrial substrate and microbial enzymes that attack this composite structure are widely used in the food, beverage, paper and pulp, and detergent sectors, while the potential utility of these enzymes in the energy industry (the estimated energy content of sugars released annually from plant cell wall degradation is equivalent to 640 billion barrels of oil) is significant. Xylan is one of the building blocks of the plant cell wall and is the major constituent of hemicellulose. After cellulose it is the most abundant renewable polysaccharide in nature. Xylan, a heterogeneous polymer and highly variable in its structure, is composed of D-xylopyranosyl units linked by β-1,4-glycosidic bonds (Fig. 1). In hardwoods, the xylan backbone is decorated with side chains, including acetic acid that esterifies the xylose units at the C-2 or C-3 positions and 4-O-methyl-D-glucuronic acid linked to the xylose units via α-1,2-glycosidic bonds. In non-acetylated softwood xylans, in addition to uronic acids, there are L-arabinofuranose residues attached to the main chain by α-1,2 and/or α-1,3-glycosidic linkages. In cereals and grasses hydroxycinnamic acids esterify the arabinofuranoses. The most abundant hydroxycinnamic acid is *trans*-ferulic acid, (*E*)-4-hydroxy-3-methoxycinnamic acid, which is usually esterified at position C-5 or C-2 to α-L-arabinofuranosyl side chains in arabinoxylans and at position C-4 to α-D-xylopyranosyl residues

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 83–97

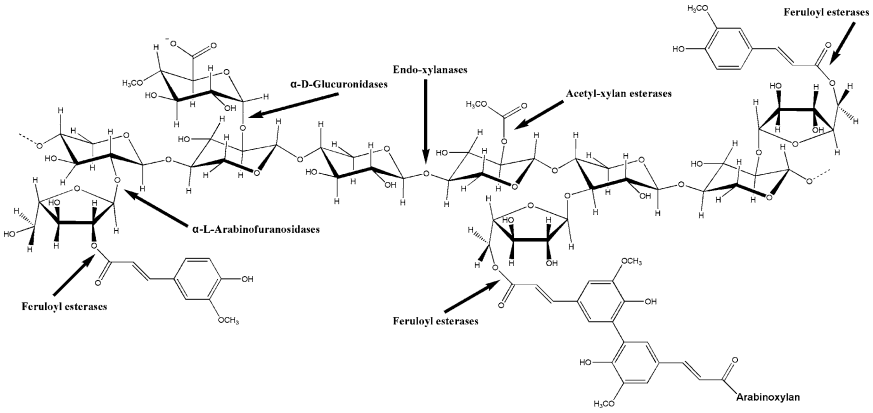


Figure 1. The basic structural components found in xylan and the hemicellulases responsible for their degradation

in xyloglucans. Chains of arabinoxylans are strengthened by cross-linking ferulic acid dimers which are ester linked to the arabinose sugars. Thus enzymes such as α-L-arabinofuranosidases (EC 3.2.1.55), α-glucuronidases (EC 3.2.1.139), acetyl xylan esterases (EC 3.1.1.72), and feruloyl esterases (EC 3.1.1.73) that remove side chain substituents from the xylan backbone are required in addition to endo-β-1,4-xylanases (EC 3.2.1.8) and β-xylosidases (EC 3.2.1.37) for the complete degradation of xylan (Fig. 1). This battery of enzymes includes two types of microbial carbohydrate esterases (EC 3.1.1.1): acetyl xylan esterases (AcXEs) and feruloyl esterases (FAEs), less well known than related lipases and other esterases.

2. ACETYL XYLAN ESTERASE

The existence of acetyl xylan esterase (AcXE, EC 3.1.1.72) was first reported in fungal cultures of *Schizophyllum commune* (Biely, 1985; Biely *et al.*, 1986). AcXE catalyses the hydrolysis of the acetyl side groups in glucuronoxylan, which is the main component of hardwood hemicellulose. Glucuronoxylan (*O*-acetyl-4-*O*-methylglucuronoxylan) is composed of β-1,4-linked D-xylopyranoside residues. Approximately every 10th xylose unit carries a 4-*O*-methylglucuronic acid side chain attached to the 2-position of xylose, and 7 out of 10 xylose residues contain an *O*-acetyl side group at the C-2 or C-3 position or both.

AcXEs fall into seven of the fourteen carbohydrate esterase (CE) families established by Coutinho and Henrissat (1999), indicating that these enzymes show considerable sequence divergence. Family 1 includes fungal AcXEs from *Aspergillus niger* (de Graaff *et al.*, 1992), *Aspergillus oryzae* (Koseki *et al.*, 2005a) other *Aspergillus* species (Koseki *et al.*, 1997; de Graaff *et al.*, 1992; Chung *et al.*, 2002; Koseki *et al.*, 2005b), *S. commune* (Halgašová *et al.*, 1994; Biely *et al.*, 1988), *Penicillium purpurogenum* AcXE I (Egaña *et al.*, 1996) and bacterial enzymes such as a Ferulic Acid Esterase (FAE) from *Cellvibrio japonicus* which contains a cellulose-binding

domain (CBD) (Ferreira *et al*., 1993), and the catalytic domains of the bifunctional enzymes of anaerobic bacteria (Fontes *et al*., 1995). Family 2 includes AcXEs from the fungus *Neocallimastix patriciarum* (Dalrymple *et al*., 1997) and the anaerobic bacterium *Clostridium thermocellum* (Hall *et al*., 1988). Family 3 includes AcXEs from the fungus *N. patriciarum* (Dalrymple *et al*., 1997), and the anaerobic rumen bacteria *C. thermocellum* (Hall *et al*., 1988) and *Ruminococcus flavefaciens* (Zhang *et al*., 1994). Family 4 includes AcXEs from *Streptomyces lividans* (contains a xylan-binding domain (XBD) (Dupont *et al*., 1996)), *Streptomyces thermoviolaceus* (Tsujibo *et al*., 1997), bifunctional enzymes (having two catalytic domains along with both CB and XB domains (Laurie *et al*., 1997; Millward-Sadler *et al.*, 1995)) from the aerobic bacteria *Cellvibrio japonicus, Cellvibrio mixtus* and *Cellulomonas fimi,* and an endoxylanase-AcXE protein from *Clostridium cellulovorans* (Kosugi *et al*., 2002). Family 5 includes fungal esterases from *Hypocrea jecorina* (formerly known as *Trichoderma reesei*) (Hakulinen *et al*., 2000) and *P. purpurogenum* (Ghosh *et al*., 2001). Family 6 includes fungal esterases from the anaerobe *N. patriciarum* (Dalrymple *et al*., 1997). Enzymes in carbohydrate esterase family 7 are unusual in that they display activity towards both acetylated xylooligosaccharides and the antibiotic cephalosporin C. Members of this family include AcXEs from *Thermoanaerobacterium* sp. (Shao *et al*., 1995), *Thermotoga maritima* (Nelson *et al*., 1999), *Bacillus pumilus*, (Krastanova *et al*., 2005) and the cephalosporin-C deacetylase from *Bacillus subtilis* (Vincent *et al*., 2003). It has been suggested that the AcXE and cephalosporin C deacetylase (EC 3.1.1.41) enzymes of the CE-7 family represent a single class of proteins with a multifunctional deacetylase activity against a range of small substrates (Vincent *et al*., 2003). Three-dimensional structures are known for AcXEs belonging to families 5 and 7. This is the case for the family 5 enzymes AXE1 of *T. reesei* (Hakulinen *et al*., 2000) and AXEII of *P. purpurogenum* (Ghosh *et al*., 2001). These two enzymes belong to the superfamily of the α/β hydrolase fold, the core domain of which has an α/β/α sandwich fold and a catalytic triad (Ser-His-Asp). Known 3-D structures of carbohydrate family 7 include *B. subtilis* cephalosporin-C deacetylase (CAH), (Vincent *et al*., 2003) and *T. maritima* AXE (Page *et al*., 2003). These enzymes are hexameric α/β hydrolases with a narrow entrance tunnel which leads to the centre of the molecule where the six active-centre catalytic triads point towards the tunnel interior and thus are sequestered away from the cytoplasmic content. By analogy to self-compartmentalising proteases, the tunnel entrance may function to hinder access of large substrates such as acetylated xylan to the poly-specific active centre. This also would explain the observation that the enzyme is active on a variety of small, acetylated molecules. The activity against cephalosporin-C suggests a possible pharmaceutical application for family 7 AcXEs in the production of semi-synthetic antibiotics.

3. FERULIC ACID ESTERASE

Ferulic Acid Esterase (FAE, EC 3.1.1.73) comprises a very diverse set of enzymes, with few sequence and physical characteristics in common. Many FAEs have been purified and characterized showing differences in physical properties such

as molecular weight, isoelectric point and optimal reaction conditions (Table 1). Multiple alignments of sequences or domains demonstrating FAE activity, as well as related sequences, have been used to construct a neighbour-joining phylogenetic tree (Crepin *et al.*, 2004a). The result of this genetic comparison, supported also by substrate specificity data, allows FAEs to be sub-classified into 4 types: A, B, C and D. Due to the increasing number of FAEs being isolated, a system of nomenclature has been proposed using the letters of the producer micro-organism followed by Fae to designate that it is an enzyme with feruloyl esterase activity and then a letter to designate the proposed sub-class based on the specificity data of the enzyme (Crepin *et al.*, 2004a). For example, the type-A FAE produced by *Fusarium oxysporum* would be termed FoFaeA. Previously reported FAEs do not follow this nomenclature. Although FAEs appear to have some common roots according to the phylogenetic tree constructed by Crepin *et al.* (2004a), they show greater sequence homology with a variety of other enzymes such as lipases, AcXEs and xylanases.

It is extremely common for esterases to act on a broad range of substrates. Esterases acting on plant cell walls catalyse similar chemical reactions but they exhibit different specificities for the aromatic moiety of hydrocinnamates or the linkage to the primary sugar in feruloylated oligosaccharides and variation in their ability to release dehydrodimeric forms of ferulic acid from plant cell wall material. The catalytic specificity shown by FAEs, as defined by the rate of catalysis divided by the Michaelis constant (k_{cat}/K_m) which gives the best indication of 'preferred' substrates, is a result of the complexity of the plant cell wall material. Type A FAEs show preference for the phenolic moiety of the substrate that contains methoxy substitutions, especially at *meta*- position(s) as occurs in ferulic and sinapinic acids, while type B FAEs show complementary activity to type A esterases, showing preference for substrates containing one or two hydroxyl substitutions as found in *p*-coumaric or caffeic acid. In contrast to type B esterases, type A FAEs appear to prefer hydrophobic substrates with bulky substituents on the benzene ring (Kroon *et al.*, 1997; Topakas *et al.*, 2005b). The high level of sequence identity of AnFaeA with the lipases from *Thermomyces lanuginosus* TLL (30% sequence identity) and *Rhizomucor miehei* (37% sequence identity) seems to justify the hydrophobic substrate preference of the esterase. Furthermore, type A and D FAEs in contrast to type B and C are also able to release small amounts of dehydrodimeric ferulic acid. Type C and D FAEs show broad specificity against synthetic hydroxycinnamic acids (ferulic, *p*-coumaric, caffeic and sinapinic acid) showing differences only in their ability to release 5-5' dehydroferulic acid (Crepin *et al.*, 2004a; Crepin *et al.*, 2004b).

Sufficient specificity studies have been conducted in order to demonstrate the ability of FAEs in releasing ferulic acid from model substrates synthesized by chemoenzymatic synthesis (Biely *et al.*, 2002) or from naturally occurring feruloylated oligosaccharides obtained by controlled enzymatic digestion of plant cell wall material. It seems that there is a correspondence between the FAE classification and the affinity of these enzymes for the position of α-L-arabinofuranose feruloylation. Type A esterases such as AnFaeA (Williamson *et al.*, 1998), TsFaeA

Table 1. Physicochemical properties of purified FAEs known to date

Micro-organism	Enzyme	FAE type	MW (kDa)	pH_{opt}	T_{opt} (°C)	pI	Reference
Aspergillus awamori	FE	-	112			3.7	McCrae *et al.*, 1994
Aspergillus awamori	CE	-	75			4.2	McCrae *et al.*, 1994
Aspergillus awamori		-	35	5.0	45	3.8	Koseki *et al.*, 1998
Aspergillus awamori	AwFAEA	A	37	5.0			Koseki *et al.*, 2005c
Aspergillus niger	FAE-I	B	63*			3.0	Faulds and Williamson, 1993
Aspergillus niger	FAE-II	A	29			3.6	Faulds and Williamson, 1993
Aspergillus niger	FAE-III or AnFaeA	A	36	5.0	55	3.3	Faulds and Williamson, 1994
Aspergillus niger	CinnAE or AnFaeB	B***	75.8*	6.0	50	4.8	Kroon *et al.*, 1996
Aspergillus niger	CE	-	120				Barbe and Dubourdieu, 1998
Aspergillus oryzae	FAE	-	30	4.5–6.0		3.6	Tenkanen *et al.*, 1991
Aspergillus tubingensis	FaeA	-	36				Vries de *et al.*, 1997
Aureobasidium pullulans		B	210	6.7	60	6.5	Rumbold *et al.*, 2003
Cellvibrio japonicus	XLYD or CjXYLD	D	59	6.0			Ferreira *et al.*, 1993
Clostridium stercorarium		C or D	33	8.0	65		Donaghy *et al.*, 2000
Clostridium thermocellum	XynZ	-	45	4–7	50–60	5.8	Blum *et al.*, 2000
Fusarium oxysporum	FoFAE-I or FoFaeB	B	31	7.0	55	>9.5	Topakas *et al.*, 2003a
Fusarium oxysporum	FAE-II or FoFaeA	A	27	7.0	45	9.9	Topakas *et al.*, 2003b
Fusarium proliferatum	FAE	B	31	6.5–7.5	50		Shin and Chen, 2005
*Lactobacillus acidophilus***		-	36	5.6	37		Wang *et al.*, 2004

(Continued)

Table 1. (Continued)

Micro-organism	Enzyme	FAE type	MW(kDa)	pH_{opt}	T_{opt} (°C)	pI	Reference
Neocallimastix MC-2	pCAE	-	11*	7.2		4.7	Borneman *et al.*, 1991
Neocallimastix MC-2	FAE-I	-	69				Borneman *et al.*, 1992
Neocallimastix MC-2	FAE-II	-	24				Borneman *et al.*, 1992
Neurospora crassa	Fae-1	B	35	6.0	55		Crepin *et al.*, 2003a
Neurospora crassa	NcFaeD-3.544	D	32				Crepin *et al.*, 2004b
Penicillium expansum		-	57.5	5.6	37		Donaghy and McKay, 1997
Penicillium funiculosum	FAE-B or PfFaeB	B	53			6.0	Kroon *et al.*, 2000
Penicillium pinophilum	p-CAE/ FAE	-	57	6.0	55	4.6	Castanares *et al.*, 1992
Piromyces equi	EstA	D	55	6.7	50–60		Fillingham *et al.*, 1999
Sporotrichum thermophile	StFAE-A or StFaeB	B	33*	6.0	55–60	3.5	Topakas *et al.*, 2004
Sporotrichum thermophile	StFaeC	C	23*	6.0	55	< 3.5	Topakas *et al.*, 2005a
Streptomyces olivochromogenes	FAE	-	29	5.5	30	7.9	Faulds and Williamson, 1991
Talaromyces stipitatus	TsFaeA	A	35			5.3	Garcia-Conesa *et al.*, 2004
Talaromyces stipitatus	TsFaeB	B	35			3.5	Garcia-Conesa *et al.*, 2004
Talaromyces stipitatus	TsFaeC	C	66	6–7	60	4.6	Crepin *et al.*, 2003b

* *Dimeric proteins (Molecular weight estimated with SDS-PAGE electrophoresis).*

** *Typical human intestinal bacterium.*

*** *Phylogenetic analysis of AnFaeB indicated that this enzyme belongs to the type C sub-class (Crepin et al., 2004a).*

(Garcia-Conesa *et al*., 2004) and FoFaeA (Topakas *et al*., 2003b) (Table 1) are active only on substrates containing ferulic acid ester linked to the C-5 and not on substrates containing ferulic acid ester linked to the C-2 linkages of L-arabinofuranose. In contrast, type B FAEs such as AnFaeB (Williamson *et al*., 1998), PfFaeB (Kroon *et al*., 2000), FAE from *A. pullulans* (Rumbold *et al*., 2003), TsFaeB (Garcia-Conesa *et al*., 2004), FoFaeB (Topakas *et al*., 2003a) and StFaeB (Topakas *et al*., 2004) (Table 1) are active on substrates containing ferulic acid ester linked to C-5 or C-2 of L-arabinofuranose, with different preferences depending on the esterase studied. The inability of type A FAE to hydrolyse the C-2 linkage between ferulic acid and the L-arabinofuranose residue could be a new criterion for use in the classification of this subclass of esterases. Type C and D FAEs such as StFaeC (Topakas *et al*., 2005a), TsFaeC (Garcia-Conesa *et al*., 2004) and CjXYLD (Ferreira *et al*., 1993) are able to hydrolyse both linkages. The active sites of FAEs from mesophilic and thermophilic sources have been probed using methyl esters of phenylalkanoic acids (Kroon *et al*., 1997; Topakas *et al*., 2005b). The thermophilic esterases from *S. thermophile* (StFaeB and StFaeC) showed k_{cat} values for phenylalkanoic and cinnamoyl methyl esters lower than those of the mesophilic esterases from *F. oxysporum* (FoFaeA and FoFaeB). A similar observation was made comparing the k_{cat} values for the methyl esters of hydroxycinnamic acids for StFaeB and FoFaeB. Lengthening or shortening the aliphatic side chain of phenylalkanoate substrates while maintaining the same aromatic substitutions of the substrates completely abolished FAE activity, showing that the distance between the aromatic group and the ester bond is critical for enzyme catalysis (Kroon *et al*., 1997; Topakas *et al*., 2005b). However, Tarbouriech *et al.* (2005) reported that substrates with short aliphatic chains (vanillate and syringate which contain only one carbon atom in the aliphatic chain) also bind to the active site of XynY FAE indicating that the length between the phenyl group and methyl ester in these molecules is not crucial, even if it may contribute to the correct orientation for catalysis.

Recently, the number of reported FAE activities has increased, especially with the acquisition of related protein sequences in genomic databases (Table 2). Many of these enzymes are modular, comprising a catalytic domain covalently fused to a non-catalytic carbohydrate binding module (Fillingham *et al*., 1999; Ferreira *et al*., 1993; Kroon *et al*., 2000; Laurie *et al*., 1997). There have also been reports of FAEs being present in large multidomain structures such as cellulosomes (Blum *et al*., 2000). A chimeric enzyme composed of feruloyl esterase A (FAEA) from *A. niger* and a dockerin from *C. thermocellum* was produced in *A. niger* (Levasseur *et al*., 2004). This is the first reported example of a functional fungal enzyme joined to a bacterial dockerin.

Unlike AcXEs which are distributed across seven different families (CE families 1 to 7), the majority of the FAEs such as the type B esterases of *N. crassa* (Crepin *et al*., 2003a) and *P. funiculosum* (Kroon *et al*., 2000) shown in Table 1 are classified in family 1 (Coutinho and Henrissat, 1999). The crystal structures of FAE, AnFaeA/FAE-III from *A. niger* (Hermoso *et al*., 2004; McAuley *et al*., 2004; Faulds *et al*., 2005) and FAE domains, XynY (Prates *et al*., 2001; Tarbouriech

Table 2. FAE genes and their accession numbers known to date

Micro-organism	Gene	Data bank	Reference
Aspergillus niger	*faeA*	Y09330	Vries de *et al.,* 1997 Juge *et al.,* 2001* Record *et al.,* 2003 Levasseur *et al.,* 2004**
Aspergillus awamori	*AwfaeA*	AB032760	Koseki *et al.,* 2005c
Aspergillus niger	*faeB*	AJ309807	Vries de *et al.,* 2002
Aspergillus tubingensis	*faeA*	Y09331	Vries de *et al.,* 1997
Butyrivibrio fibrisolvents	*cinI* or*cinA*	U44893	Dalrymple *et al.,* 1996
Butyrivibrio fibrisolvents	*cinII* or*cinB*	U64802	Dalrymple and Swadling, 1997
Cellvibrio japonicus	*xynD*	X58956	Ferreira *et al.,* 1993
Clostridium thermocellum	*XynY*	X83269	Blum *et al.,* 2000
Clostridium thermocellum	*XynZ*	M22624	Blum *et al.,* 2000
Neurospora crassa	*Fae-1*	AJ293029	Crepin *et al.,* 2003a*
Neurospora crassa	*faeD-3.544*	-	Crepin *et al.,* 2004b
Penicillium funiculosum	*faeB*	AJ291496	Kroon *et al.,* 2000
Piromyces equi	*estA*	AF164516	Fillingham *et al.,* 1999
Talaromyces stipitatus	*faeC*	AJ505939	Crepin *et al.,* 2003b*

* *Heterologous expression of the FAE in the methylotrophic yeast Pichia pastoris.*
** *Chimeric protein associating FAEA from A. niger and a dockerin domain from C. thermocellum.*

et al., 2005), and XynZ (Schubot *et al*, 2001) of the cellulosomal enzymes included in the cellulosome complex from *C. thermocellum*, have been determined. These FAEs have a common α/β hydrolase fold and a catalytic triad (Ser-His-Asp) also present in lipases. For example, the structure of AnFaeA displays an α/β hydrolase fold very similar to that of the fungal lipases from *T. lanuginosus* (Lawson *et al*, 1994) and *R. miehei* (Derewenda *et al*, 1992) but lacks lipase activity (Aliwan *et al*, 1999). The active site cavity is confined by a lid, similar to that of lipases, and by a loop that confers plasticity to the substrate binding site. The lid presents a high ratio of polar residues, which, in addition to a unique *N*-glycosylation site, stabilizes the lid in an open conformation conferring the esterase character to this enzyme (Hermoso *et al*, 2004). The structure and the sequence homology of AnFaeA are different from that reported for the cellulosomal enzymes XynY and XynZ from *C. thermocellum*, although the catalytic triads can be superimposed allowing direct extrapolation of the position of the oxyanion pocket. Co-crystallization studies of the inactive forms of XynY and AnFaeA with ferulic acid (Prates *et al*, 2001; McAuley *et al*, 2004) or XynZ and AnFaeA with feruloyl oligosaccharides (Schubot *et al*, 2001; Faulds *et al*, 2005) were conducted in order to identify the residues involved in substrate binding and reveal the hydrolytic mechanism of FAEs. Furthermore, the structures of XynY FAE Ser-Ala mutant complexes with syringate, sinapinate and vanillate methyl esters were reported by Tarbouriech *et al* (2005) indicating the importance of the *meta*-methyl group of the ferulic ring for binding. Faulds *et al* (2005) solved the crystal structure of

an inactive mutant of AnFaeA (S133A) in complex with *O*-{5-*O*-[(E)-feruloyl]-α-L-arabinofuranosyl}-(1→3)-*O*-β-D-xylopyranosyl-(1→4)-D-xylopyranose (FAX_2) and observed that the ferulic acid moiety of the substrate was visible in the electron density map showing interactions through its OH and OCH_3 groups with the hydroxyl groups of Tyr80. However, the remaining groups of the substrate (*i.e.* the arabinose and the two xylose units) were not visible. Accordingly, in the structure of the XynZ FAE in complex with FAX_2 determined by Schubot *et al.* (2001), the ferulic moiety was clearly visible in the active site while the carbohydrate parts of the substrate were not, suggesting that tight binding of the carbohydrate is not required for catalysis. These results are in agreement with the synthetic ability of StFaeC in non-conventional media where the esterase seems to be able to esterify a broad spectrum of sugars showing specificity only to the ferulic moiety (Vafiadi *et al.*, 2005). In contrast to FAEs, determination of the crystal structure of a family 10 xylanase from *Thermoascus aurantiacus* complexed with xylobiose containing an arabinofuranosyl-ferulate side-chain, revealed that the distal glycone subsite of the enzyme makes extensive direct and indirect interactions with the arabinose side-chain, while the ferulate moiety is solvent-exposed (Vardakou *et al.*, 2005).

4. USE OF ACETYL XYLAN ESTERASES AND FERULIC ACID ESTERASES AS BIOSYNTHETIC TOOLS

There is an increasing demand for "green" (environmentally friendly) production processes for biodegradable polymers with modified hydrophobic and rheological characteristics. Enzymatic acylation of oligo- and polysaccharides is more environmentally friendly than classical chemical synthesis. AcXE from *S. commune*, a member of CE family 1 catalyses acetyl group transfer to methyl β-D-xylopyranoside and other substrates (Biely *et al.*, 2003). This work was the first published example of reverse reactions by AcXE.

Various hydroxycinnamic acids (ferulic, *p*-coumaric, caffeic, sinapinic) have widespread industrial potential by virtue of their antioxidant properties. Generally, such natural antioxidants are partially soluble in aqueous media, limiting their usefulness in oil-based processes and that has been reported to be a serious disadvantage if an aqueous phase is also present. The modification of these compounds via esterification with aliphatic alcohols results in the formation of more lipophilic derivatives. The direct esterification of natural phenolic acids including the above mentioned hydroxycinnamic acids with aliphatic alcohols catalysed by various lipases in organic media has been reported, albeit with low reaction rate and yield. Several authors have demonstrated that the lipase-inhibiting effect of electron-donating substituents conjugated to the carboxylic groups in hydroxylated derivatives of cinnamic acids like ferulic, *p*-coumaric, sinapinic and caffeic acid, is strong (Figueroa-Espinoza and Villeneuve, 2005). Esterification can be carried out by lipases only if the aromatic ring is not *para*-hydroxylated and the lateral chain is saturated. Thus, the enzymatic esterification of cinnamoyl substrates can be obtained using only FAEs as biocatalysts.

Transesterification of phenolic acids was catalysed by using a type A FAE from *F. oxysporum* (FoFaeA) trapped in a *n*-hexane/1-propanol/water surfactantless microemulsion (Topakas *et al*., 2003a). Greater synthetic activity was observed in ternary water-organic mixtures having a lower water content. The synthetic activity of esterases follows a pattern similar to their hydrolytic activity against various methyl esters of cinnamic acids. FoFaeA shows a preference for the hydrolysis of methoxylated substrates (Topakas *et al*., 2005b) while conversion to butyl esters was greater with ferulic and sinapinic acids. Type B esterases from *F. oxysporum* (FoFaeB) (Topakas *et al*., 2003b) and *S. thermophile* (StFaeB) showed preference for the hydrolysis of hydroxylated substrates (Topakas *et al*., 2005b) and the conversion to butyl esters was enhanced with *p*-coumaric and caffeic acids (Topakas *et al.*, 2003b, 2004). The type-C FAE StFaeC from *S. thermophile* demonstrated maximum hydrolytic activity against methyl ferulate (Topakas *et al*., 2005b). Optimal yields were achieved producing butyl esters with ferulic acid (Topakas *et al.*, 2005a). Furthermore, it was reported that the same enzyme catalysed the transfer of the feruloyl group to L-arabinose (Fig. 2) in a ternary water-organic mixture consisting of *n*-hexane, *t*-butanol and water system, achieving a conversion of about 40% of L-arabinose to a feruloylated derivative (Topakas *et al*., 2005a). This work was the first example of sugar esterification with unsaturated arylaliphatic acids, like methoxylated or hydroxylated derivatives of cinnamic acids (such as ferulic acid). Lipases are not able to catalyse such a reaction due to electronic and/or steric effects (Otto *et al*., 2000).

Phenolic acid sugar esters have demonstrable antitumoural activity and the potential to be used to formulate antimicrobial, antiviral and/or anti-inflammatory agents. As esters based on unsaturated arylaliphatic acids such as cinnamic acid and its derivatives are known to display anticancer activity, specific FAEs could be employed in the tailored synthesis of such pharmaceuticals.

The potential use of FAEs for the synthesis of feruloylated oligomers or polymers using feruloyl esterases opens the door for the design of modified biopolymers with new properties and bioactivities.

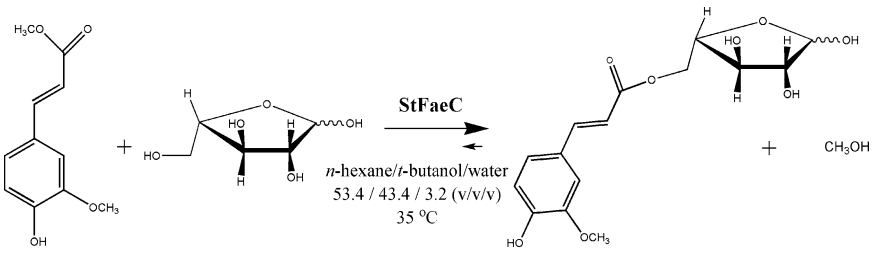


Figure 2. Transesterification of methyl ferulate with L-arabinose by StFaeC

REFERENCES

Aliwan, F.O., Kroon, P.A., Faulds, C.B., Pickersgill, R. and Williamson, G. (1999) Ferulic acid esterase-III from *Aspergillus niger* does not exhibit lipase activity. J. Sci. Food Agric. *79*, 457–459.

Barbe, C. and Dubourdieu, D. (1998) Characterization and purification of a cinnamate esterase from *Aspergillus niger* industrial pectinase preparation. J. Sci. Food Agric. *78*, 471–478.

Biely, P. (1985) Microbial xylanolytic system. Trends Biotechnol. *3*, 286–290.

Biely, P., MacKenzie, C.R., Puls, J. and Schneider, H. (1986) Cooperativity of esterases and xylanases in the enzymatic degradation of acetyl xylan. Bio Technol. *4*, 731–733.

Biely, P., MacKenzie, C.R. and Schneider, H. (1988) Production of acetyl xylan esterase by *Trichoderma reesei* and *Schizophyllum commune*. Can. J. Microbiol. *34*, 767–772.

Biely, P., Mastihubova, M., van Zyl, W.H. and Prior, B.A. (2002) Differentiation of feruloyl esterases on synthetic substrates in alpha-arabinofuranosidase-coupled and UV-spectrophotometric assays. Anal. Biochem. *311*, 68–75.

Biely, P., Wong, K.K.Y., Suckling, I.D. and Spanikova, S. (2003) Transacetylations to carbohydrates catalyzed by acetyl xylan esterase in the presence of organic solvent. Biochem. Biophys. Acta *1623*, 62–71.

Blum, D.L., Kataeva, I.A., Li, X.L. and Ljungdahl, L.G. (2000) Feruloyl esterase activity of the *Clostridium thermocellum* cellulosome can be attributed to previously unknown domains of XynY and XynZ. J. Bacteriol. *182*, 1346–1351.

Borneman, W.S., Ljungdahl, L.G., Hartley, R.D. and Akin, D.E. (1991) Isolation and characterization of p-coumaroyl esterase from the anaerobic fungus *Neocallimastix* strain MC-2. Appl. Environ. Microbiol. *57*, 2337–2344.

Borneman, W.S., Ljungdahl, L.G., Hartley, R.D. and Akin, D.E. (1992) Purification and partial characterisation of two feruloyl esterases from the anaerobic fungus *Neocallimastix* strain MC-2. Appl. Environ. Microbiol. *58*, 3762–3766.

Castanares, A., McCrae, S.I. and Wood, T.M. (1992) Purification and properties of a feruloyl/p-coumaroyl esterase from the fungus *Penicillium pinophilum*. Enzyme Microb. Technol. *14*, 875–884.

Chung, H.-J., Park, S.-M., Kim, H.R., Yang, M.S. and Kim, D.H. (2002) Cloning the gene encoding acetyl xylan esterase from *Aspergillus ficuum* and its expression in *Pichia pastoris*. Enzyme Microb. Technol. *31*, 384–391.

Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-active enzymes, http:// afmb.cnrs-mrs.fr/CAZY/, server at URL.

Crepin, V.F., Faulds, C.B. and Connerton, I.F. (2003a) A non-modular type B feruloyl esterase from *Neurospora crassa* exhibits concentration-dependent substrate inhibition. Biochem. J. *370*, 417–427.

Crepin, V.F., Faulds, C.B. and Connerton, I.F. (2003b) Production and characterization of the *Talaromyces stipitatus* feruloyl esterase FAEC in *Pichia pastoris*: identification of the nucleophilic serine. Protein Expres. Purif. *29*, 176–184.

Crepin, V.F., Faulds, C.B. and Connerton, I.F. (2004a) Functional classification of the microbial feruloyl esterases. Appl. Microbiol Biotechnol. *63*, 647–652.

Crepin, V.F., Faulds, C.B. and Connerton, I.F. (2004b) Identification of a type-D feruloyl esterase from *Neurospora crassa*. Appl. Microbiol. Biotechnol. *63*, 567–570.

Dalrymple, B.P., Swadling, Y., Cybinski, D.H. and Xue, G.-P.(1996) Cloning of a gene encoding cinnamoyl ester hydrolase from the ruminal bacterium *Butyrivibrio fibrisolvens* E14 by a novel method. FEMS Microbiol. Lett. *143*, 115–120.

Dalrymple, B.P. and Swadling, Y. (1997) Expression of a *Butyrivibrio fibrisolvens* E14 gene *(cinB)* encoding an enzyme with cinnamoyl esterase activity is negatively regulated by the product of an adjacent gene *(cinR).* Microbiology *143*, 1203–1210.

Dalrymple, B.P., Cybinski, D.H., Layton, I., McSweeney, C.S., Xue, G.-P, Swadling, Y.J. and Lowry J.B. (1997) Three *Neocallimastix patriciarum* esterases associated with the degradation of complex polysaccharides are members of a new family of hydrolases. Microbiology *143*, 2605–2614.

Derewenda, Z.S., Derewenda, U. and Dodson, G. (1992) The crystal and molecular structure of the *Rhizomucor miehei* triacylglyceride lipase at 1.9 Å resolution. J. Mol. Biol. *227*, 818–839.

Donaghy, J. and McKay, A.M. (1997) Purification and characterization of a feruloyl esterase from the fungus *Penicillium expansum*. J. Appl. Microbiol. *83*, 718–726.

Donaghy, J.A., Bronnenmeier, K., Soto-Kelly, P.F. and McKay, A.M. (2000) Purification and characterization of an extracellular feruloyl esterase from the thermophilic anaerobe *Clostridium stercorarium*. J. Appl. Microbiol. *88*, 458–466.

Dupont, C., Daigneault, N., Shareck, F., Morosoli, R. and Kluepfel, D. (1996) Purification and characterization of an acetyl xylan esterase produced by *Streptomyces lividans*. Biochem. J. *319*, 881–886.

Egaña, L., Gutiérrez, R., Caputo, V., Peirano, A., Steiner, J. and Eyzaguirre, J. (1996) Purification and characterization of two acetyl xylan esterases from *Penicillium pupurogenum*. Biotechnol. Appl. Biochem. *24*, 33–39.

Faulds, C.B. and Williamson, G. (1991) The purification and characterisation of 4-hydroxy-3methoxycinnamic (ferulic) acid esterase from *Streptomyces olivochromogenes*. J. Gen. Microbiol. *137*, 2339–2345.

Faulds, C.B. and Williamson, G.(1993) Ferulic acid esterase from *Aspergillus niger:* purification and partial characterization of two forms from a commercial source of pectinase. Biotech. Appl. Biochem. *17*, 349–359.

Faulds, C.B. and Williamson, G. (1994) Purification and characterization of a ferulic acid esterase (FAE-III) from *Aspergillus niger*: specificity for the phenolic moiety and binding to microcrystalline cellulose. Microbiology *140*, 779–787.

Faulds, C.B., Molina, R., Gonzalez, R., Husband, F., Juge, N., Sanz-Aparicio, J. and Hermoso, J.M. (2005) Probing the determinants of substrate specificity of a feruloyl esterase, AnFaeA, from *Aspergillus niger*. FEBS J. *272*, 4362–4371.

Ferreira, L.M.A., Wood, T.M., Williamson, G., Faulds, C.B., Hazlewood, G.P. and Gilbert, H.J. (1993) A modular esterase from *Pseudomonas fluorescens* subsp. *cellulosa* contains a non-catalytic binding domain. Biochem. J. *294*, 349–355.

Figueroa-Espinoza, M.-C. and Villeneuve, P. (2005) Phenolic acids enzymatic lipophilization. J. Agric. Food Chem. *53*, 2779–2787.

Fillingham, I.J., Kroon, P.A., Williamson, G., Gilbert, H.J. and Hazlewood, G.P. (1999) A modular cinnamoyl ester hydrolase from the anaerobic fungus *Piromyces equi* acts synergistically with xylanase and is part of a multiprotein cellulose-binding cellulose-hemicellulase complex. Biochem. J. *343*, 215–224.

Fontes, C.M., Hazlewood, G.P., Morag, E., Hall, J., Hirst, B.H. and Gilbert., H.J. (1995) Evidence for a general role for non-catalytic thermostabilizing domains in xylanases from thermophilic bacteria. Biochem. J. *307*, 151–158.

Garcia-Conesa, M.T., Valerie, F.C., Goldson, A.J., Williamson, G., Cummings, N.J., Connerton, I.F., Faulds, C.B. and Kroon, P.A. (2004) The feruloyl esterase system of *Talaromyces stipitatus*: production of three discrete feruloyl esterases, including a novel enzyme, TsFaeC, with a broad substrate specificity. J. Biotechnol. *108*, 227–241.

Ghosh, D., Sawichi, M., Lala, P., Erman, M., Pangborn, W., Eyzaguirre, J., Gutiérrez, R., Jörnvall, H. and Thiel, D.J. (2001) Multiple conformation of catalytic serine and histidine in acetyl xylan esterase at 0.90 Å. J. Biol. Chem. *276*, 11159–11166.

de Graaff, L.H., Visser, J., van den Broeck, H.C., Strozyk, F., Kormelink, F.J.M. and Boonman, J.C.P. (1992) Cloning, expression and use of acetyl xylan esterases from fungal origin. EP 0507369-A 7 07-OCT-1992

Hakulinen, N., Tenkanen, M. and Rouvinen, J. (2000) Three-dimensional structure of the catalytic core of acetyl xylan esterase from *Trichoderma reesei*: insights into the deacetylation mechanism. J. Struct. Biol. *132*, 180–190.

Halgašová, N., Kutejová, E. and Timko, J. (1994) Purification and some characteristics of the acetyl xylan esterase from *Schizophyllum commune*. Biochem. J. *298*, 751–755.

Hall, J., Hazlewood, G.P., Baker, P.J. and Gilbert, H.J. (1988) Conserved reiterated domains in *Clostridium thermocellum* endoglucanases are not essential for catalytic activity. Gene *69*, 29–38.

Hermoso, J.A., Sanz-Aparicio, J., Molina, R., Juge, N., Gonzalez, R. and Faulds, C.B. (2004) The crystal structure of feruloyl esterase A from *Aspergillus niger* suggests evolutive functional convergence in feruloyl esterase family. J. Mol. Biol. *338*, 495–506.

Juge, N., Williamson, G., Puigserver, A., Cummings, N.J., Connerton, I.F. and Faulds, C.B. (2001) High-level production of recombinant *Aspergillus niger* cinnamoyl esterase (FAEA) in the methylotrophic yeast *Pichia pastoris*. FEMS Yeast Res. *1*, 127–132.

Koseki, T., Furuse, S., Iwano, K., Sakai, H. and Matsuzawa, H. (1997) An *Aspergillus awamori* acetylesterase: purification of the enzyme, and cloning and sequencing of the gene. Biochem. J. *326*, 485–490.

Koseki, T., Furuse, S., Iwano, K. and Matsuzawa, H. (1998) Purification and characterization of a feruloylesterase from *Aspergillus awamori*. Biosci. Biotech. Biochem. *62*, 2032–2034.

Koseki, T., Miwa, Y., Akao, T., Akita, O. and Hashizume, K. (2005a) An *Aspergillus oryzae* acetyl xylan esterase: Molecular cloning and characteristics of recombinant enzyme expressed in *Pichia pastoris* J. Biotechnol. in press

Koseki, T., Miwa, Y., Fushinobu, S. and Hashizume, K. (2005b) Biochemical characterization of recombinant acetyl xylan esterase from *Aspergillus awamori* expressed in *Pichia pastoris*: Mutational analysis of catalytic residues. Biochim. Biophy. Acta *1749*, 7–13.

Koseki, T., Takahashi, K., Fushinobu, S., Iefuji, H., Iwano, K., Hashizume, K. and Matzuzawa, H. (2005c) Mutational analysis of a feruloyl esterase from *Aspergillus awamori* involved in substrate discrimination and pH dependence. Biochem. Biophys. Acta *1722*, 200–208.

Kosugi, A., Murashima, K. and Doi, R.H. (2002) Xylanase and acetyl xylan esterase activities of XynA, a key subunit of the *Clostridium cellulovorans* cellulosome for xylan degradation. Appl. Environ. Microbiol. *68*, 6399–6402.

Krastanova, I., Guarnaccia, C., Zahariev, S., Degrassi, G. and Lamba, D. (2005) Heterologous expression, purification, crystallization, X-ray analysis and phasing of the acetyl xylan esterase from *Bacillus pumilus.* Biochem. Biophys. Acta *1748*, 222–230.

Kroon, P.A., Faulds, C.B. and Williamson, G. (1996) Purification and characterisation of a novel esterase induced by growth of *Aspergillus niger* on sugar-beet pulp. Biotechnol. Appl. Biochem. *23*, 255–262.

Kroon, P.A., Faulds, C.B., Brezillon, C. and Williamson, G. (1997) Methyl phenylalkanoates as substrates to probe the active sites of esterases. Eur. J. Biochem. *248*, 245–251.

Kroon, P.A., Williamson, G., Fish, N.M., Archer, D.B. and Belshaw, N.J. (2000) A modular esterase from *Penicillium funiculosum* which releases ferulic acid from plan cell walls and binds crystalline cellulose contains a carbohydrate binding module. Eur. J. Biochem. *267*, 6740–6752.

Laurie, J.I., Clarke, J.H., Ciruela, A., Faulds, C.B., Williamson, G., Gilbert, H.J., Rixon, J.E., Millward-Sadler, J. and Hazlewood, G.P. (1997) The NodB domain of a multifunctional xylanase from *Cellulomonas fimi* deacetylates acetyl xylan. FEMS Microbiol. Lett. *148*, 261–264.

Lawson, D.M., Brzozowski, A.M., Dodson, G.G., Hubbard, R.E., Huge-Jensen, B., Boel, E. and Derewenda, Z.S. (1994) The three-dimensional structures of two lipases from filamentous fungi. In Lipases: Their Biochemistry, Structure and Application (Woolley P and Petersen S, eds), pp. 77–94. Cambridge University Press, Cambridge, UK.

Levasseur, A., Pages, S., Fierobe, H.-P., Navarro, D., Punt, P., Belaich, J.-P., Asther, M. and Record, E. (2004) Design and production in *Aspergillus niger* of a chimeric protein associating a fungal feruloyl esterase and a clostridial dockerin domain. Appl. Environ. Microbiol. *70*, 6984–6991.

McAuley, K., Svendsen, A., Patkar, S.A. and Wilson, K.S. (2004) Structure of a feruloyl esterase from *Aspergillus niger*. Acta Cryst. D. *60*, 878–887.

McCrae, S.I., Leith, K.M., Gordon, A.H. and Wood, T.M. (1994) Xylan degrading enzyme system produced by the fungus *Aspergillus awamori*: isolation and characterization of a feruloyl esterase and a p-coumaroyl esterase. Enzyme Microb. Technol. *16*, 826–834.

Millward-Sadler, S.J., Davidson, K., Hazlewood, G.P., Black, G.W., Gilbert, H.J. and Clarke, J.H. (1995) Novel cellulose-binding domains, NodB homologous and conserved modular architecture in xylanases from the aerobic soil bacteria *Pseudomonas fluorescens* subsp. *cellulosa* and *Cellvibrio mixtus*. Biochem. J. *312*, 39–48.

Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L., Dodson, R.J., Haft, D.H. *et al.* (1999) Evidence for lateral gene transfer between archaea and bacteria from genome sequence of *Thermatoga maritima*. Nature *399*, 323–329.

Otto, R.T., Scheib, H., Bornscheuer, U.T., Pleiss, J., Syldatk, C. and Schmid, R.D. (2000) Substrate specificity of lipase B from *Candida antarctica* in the synthesis of arylaliphatic glycolipids. J. Mol. Catal. B: Enzymatic. *8*, 201–211.

Page, R., Crzechnik, S.K., Canaves, J.M., Spraggon, G., Kreusch, A., Kuhn, P., Stevens, R.C. and Lesley, S.A. (2003) Shotgun crystallization strategy for structural genomics: an optimized two-tiered crystallization screen against the *Thermotoga maritima* proteome. Acta Crystallogr. Sect. D Biol. Crystallogr. *59*, 1028–1037.

Prates, J.A.M., Tarbouriech, N., Charnock, S.J., Fontes, C.M.G.A., Ferreira, L.M.A. and Davies, G.J. (2001) The structure of the feruloyl esterase module of xylanase 10B from *Clostridium thermocellum* provides insights into substrate recognition. Structure *9*, 1183–1190.

Record, E., Asther, M., Sigoillot, C., Pages, S., Punt, P.J., Delattre, M., Haon, M., van del Hondel, C.A.M.J.J., Sigoillot, J.-C., Lesage-Meessen, L. and Asther, M. (2003) Overproduction of the *Aspergillus niger* feruloyl esterase for pulp bleaching application. Appl. Microbiol. Biotechnol. *62*, 349–355.

Rumbold, K., Biely, P., Mastihubova, M., Gudelj, M., Gubitz, G., Robra, K.-H. and Prior, B.A. (2003) Purification and properties of a feruloyl esterase involved in lignocellulose degradation by *Aureobasidium pullulans*. Appl. Environ. Microbiol. *69*, 5622–5626.

Schubot, F.D., Kataeva, I.A., Blum, D.L., Shah, A.K., Ljungdahl, L.G., Rose, J.P. and Wang, B.C. (2001) Structural basis for the substrate specificity of the feruloyl esterase domain of the cellulosomal xylanase Z from *Clostridium thermocellum.* Biochemistry *40*, 12524–12532.

Shao, W. and Wiegel, J. (1995) Purification and characterization of two thermostable acetyl xylan esterases from *Thermoanaerobacterium* sp. strain JW/SL YS485. Appl. Environ. Microbiol. *61*, 729–733.

Shin, H-D. and Chen, R.R. (2006) Production and characterization of a type B feruloyl esterase from *Fusarium proliferatum* NRRL 26517. Enzyme Microb. Technol. *38*, 478–485.

Tarbouriech, N., Prates, J.A.M., Fontes, C.M.G.A. and Davies, G.J. (2005) Molecular determinants of substrate specificity in the feruloyl esterase module of xylanase 10B from *Clostridium thermocellum*. Acta Cryst. D. *61*, 194–197.

Tenkanen, M., Schuseil, J., Puls, J. and Poutanen, K. (1991) Production, purification and characterisation of an esterase liberating phenolic acids from lignocellulose. J. Biotechnol. *18*, 69–84.

Topakas, E., Stamatis, H., Mastihubova, M., Biely, P., Kekos, D., Macris, B.J, and Christakopoulos, P. (2003a) Purification and characterization of a *Fusarium oxysporum* feruloyl esterase (FoFAE-I) catalysing transesterification of phenolic acid esters. Enzyme Microb. Technol. *33*, 729–737.

Topakas, E., Stamatis, H., Biely, P., Kekos, D., Macris, B.J. and Christakopoulos, P. (2003b) Purification and characterization of a feruloyl esterase from *Fusarium oxysporum* catalyzing esterification of phenolic acids in ternary water-organic solvent mixtures. J. Biotech. *102*, 33–44.

Topakas, E., Stamatis, H., Biely, P. and Christakopoulos, P. (2004) Purification and characterization of a type B feruloyl esterase (StFAE-A) from the thermophilic fungus *Sporotrichum thermophile*. Appl. Microbiol. Biotechnol. *63*, 686–690.

Topakas, E., Vafiadi, C., Stamatis, H. and Christakopoulos, P. (2005a) *Sporotrichum thermophile* type C feruloyl esterase (StFaeC): Purification characterization, and its use for phenolic acid (sugar) ester synthesis. Enzyme Microb. Technol. *36*, 729–736.

Topakas, E., Christakopoulos, P. and Faulds, C.B. (2005b) Comparison of mesophilic and thermophilic feruloyl esterases: characterization of their substrate specificity for methyl phenylalkanoates. J. Biotechnol. *115*, 355–366.

Tsujibo, H., Ohtsuki, T., Ito, T., Yamazaki, I., Miyamoto, K., Sugiyama, M. and Inamori, Y. (1997) Cloning and sequence analysis of genes encoding xylanases and acetyl xylan esterase from *Stretomyces thermoviolaceus* OPC-520. Appl. Environ. Microb. *63*, 661–664.

Vardakou, M., Flint, J., Christakopoulos, P., Lewis, R.J., Gilbert, H.J. and Murray, J.W. (2005) A family 10 *Thermoascus aurantiacus* xylanase utilizes arabinose decorations of xylan as significant substrate specificity determinants. J. Mol. Biol. *352*, 1060–1067.

Vafiadi, C., Topakas, E., Wong, K.K.Y., Suckling, I.D. and Christakopoulos, P. (2005) Mapping the hydrolytic and synthetic selectivity of a type C feruloyl esterase (StFaeC) from *Sporotrichum thermophile* using alkyl ferulates. Tetrahedron Asymmetry *16*, 373–379.

Vincent, F., Charnock, S.J., Verschueren, K.H.G., Turkenburg, J.P., Scott, D.J., Offen, W.A., Roberts, S., Pell, G., Gilbert, H.J., Davies, G.J. and Brannigam, J.A. (2003) Multifunctional xylooligosaccharide/cephalosporin C deacetylase revealed by the hexameric structure of the *Bacilus subtilis* enzyme at 1.9 Å resolution. J. Mol. Biol. *330*, 593–606.

de Vries, R.P., Michelsen, B., Poulsen, C.H., Kroon, P.A., Heuvel van den, R.H.H., Faulds, C.B., Williamson, G., Hombergh van den, J.P.T.W. and Visser, J. (1997) The faeA genes from *Aspergillus niger* and *Aspergillus tubingensis* encode ferulic acid esterases involved in degradation of complex cell wall polysaccharides. Appl. Environ. Microbiol. *63*, 4638–4644.

de Vries, R.P., vanKuyk, P.A., Kester, H.C.M. and Visser, J. (2002) The *Aspergillus niger faeB* gene encodes a second feruloyl esterase involved in pectin and xylan degradation and is specifically induced in the presence of aromatic compounds. Biochem. J. *363*, 377–386.

Wang, X., Geng, X., Egashira, Y. and Sanada, H. (2004) Purification and characterization of a feruloyl esterase from the intestinal bacterium *Lactobacillus acidophilus.* Appl. Environ. Microbiol. *70*, 2367–2372.

Williamson, G., Faulds, C.B. and Kroon, P.A. (1998) Specificity of ferulic acid (feruloyl) esterases. Biochem. Soc. Trans. *26*, 205–209.

Zhang, J.-X., Martin, J. and Flint, H.J. (1994) Identification of non-catalytic conserved regions in xylanases encoded by the xynB and xynD genes of the cellulolytic rumen anaerobe *Ruminococcus flavefaciens*. Mol. Gen. Genet. *254*, 260–264.

CHAPTER 7

STRUCTURAL AND BIOCHEMICAL PROPERTIES OF PECTINASES

SATHYANARAYANA N. GUMMADI*, N. MANOJ AND D. SUNIL KUMAR
Department of Biotechnology, Indian Institute of Technology-Madras, Chennai, India
**gummadi@iitm.ac.in*

1. INTRODUCTION

Pectin and other pectic substances are complex polysaccharides, which contribute firmness and structure to plant tissues as a part of the middle lamella. The basic unit in pectic substances is galacturonan (α-D-galacturonic acid). Pectic substances are classified into two types; homogalacturonan and heterogalacturonan (rhamnogalacturonan). In homogalacturonan, the main polymer chain consists of α-D-galacturonate units linked by $(1 \rightarrow 4)$ glycosidic bonds, whereas in rhamnogalacturonan, the primary chain consist of $(1 \rightarrow 4)$ linked α-D-glacturonates and with about 2–4% L-rhamnose units that are $\beta(1 \rightarrow 2)$ and $\beta(1 \rightarrow 4)$ linked to D-galacturonate units (Whitaker, 1991). The side chains of rhamnogalacturonans usually consist of L-arabinose or D-galacturonic acid units. In plant tissues, about 60–70% of the galacturonate units are esterified with methanol and occasionally with ethanol. Based on the degree of esterification, pectic substances are classified into protopectin, pectinic acid, pectin and polygalacturonic acid (Table 1). Molecular size, degree of esterification and weight distribution of polygalacturonic acid residues are important factors that contribute to heterogeneity in pectic substances. Relative molecular masses of pectic substances isolated from various sources such as citrus fruits, apple and plums, range from 25 to 350 kDa.

Pectinases are a complex and diverse group of enzymes involved in the degradation of pectic substances. The diversity of forms of pectic substances in plant cells probably accounts for the existence of various forms of these enzymes. Pectinases are classified depending on their substrate and mode of enzymatic reaction (Fig. 1). Pectinases act as carbon recycling agents in nature by degrading pectic substances to saturated and unsaturated galacturonans, which are further catabolized

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 99-115

Table 1. Different types of pectic substances[a]

Pectic substances	Structural description	Properties
Protopectin	Galacturonate units linked by α-1,4-glycosidic linkages. The carboxyl groups are highly esterified with methanol. Polymer is highly cross-linked with Ca^{2+} or with other polysaccharides	Insoluble in water. Degree of esterification > 90%
Pectin	Galacturonate units linked by α-1,4-glycosidic linkages. The carboxyl groups are esterified with methanol	Soluble in water. Degree of esterification at least 75%
Pectinic acid	Galacturonate units linked by α-1,4-glycosidic linkages. The carboxyl groups are slightly esterified with methanol	Soluble in water. Degree of esterification varies between 0 and 75%
Pectic acid	Galacturonate units linked by α-1,4-glycosidic linkages.	Soluble in water. Degree of esterification 0
Rhamnogalacturonan	Galacturonate units linked by α-1,4-glycosidic linkages with rhamnose units lined by β-1,2 and β-1,4 linkages. The side chains are homogeneous polymers of galacturonic acid and arabinose	Soluble in water.

[a]Apart from those mentioned in the table, oligomers of galacturonate and methyl galacturonate are present.

to 5-keto-4-deoxy-uronate and finally to pyruvate and 3-phosphoglyceraldehyde (Vincent-Sealy *et al*, 1999). Pectinases from phytopathogenic fungi such as *Aspergillus flavus*, *Fusarium oxysporum* and *Botrytis cinerea* are also known to play a vital role in plant pathogenicity or virulence by degrading pectic compounds present in cell wall (Lang and Dörenberg, 2000; DiPietro and Roncero, 1996; Ten Have *et al.*, 1998). Pectinases, especially polygalacturonase, is known to play a major role in pectin breakdown during the final stages of fruit ripening (Sozzi-Quiroga and Fraschina, 1997; Chin *et al*, 1999). Polymethylgalacturonase (PMG), polygalacturonase (PG), pectin lyase (PL), polygalacturonate lyase (PGL) and pectinesterase (PME) are industrially important pectinases discussed in this chapter.

2. BIOCHEMICAL CHARACTERISTICS OF PECTINASES

2.1. Polymethylgalacturonase

PMG activity can be determined by measuring the reducing sugars formed due to the hydrolysis of glycosidic bond or by measuring the reduction in viscosity of the substrate. Highly esterified pectin is the best substrate for PMG whereas pectic acid and pectate derivatives do not react with PMG. Few reports are available on the biochemical characteristics of the enzyme due to following reasons: (a) the enzyme has not been purified to homogeneity and characterized and (b) the activity of this enzyme has not been demonstrated in the absence of other pectic enzymes and in the presence of 100% methylated pectin. Hence, researchers need to be very careful in reporting the activity of PMG. *Aspergillus* was found to be a major producer

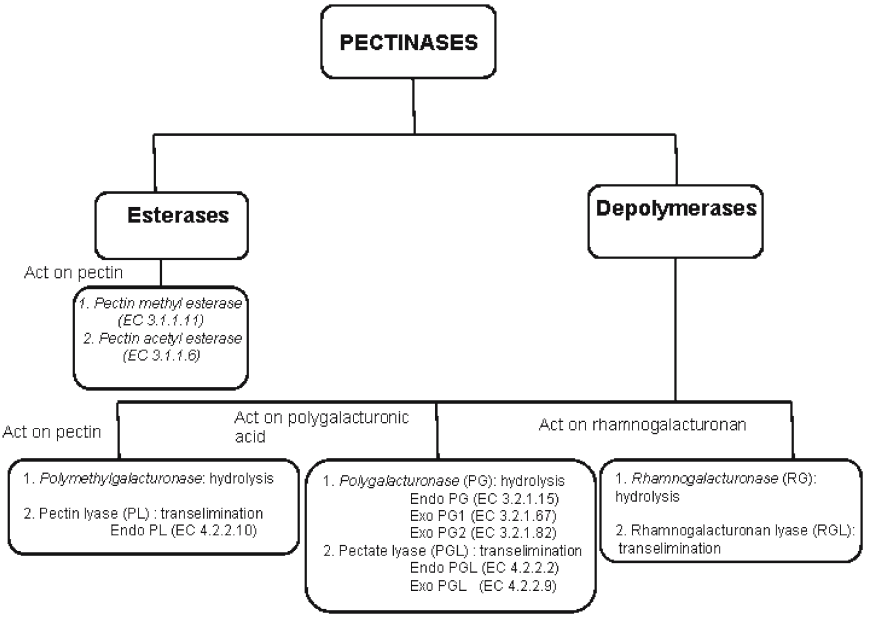


Figure 1. Classification of different pectinases based on their reaction with different pectic substances

of PMG followed by species belonging to *Penicillium*, *Botrytis* and *Sclerotium*. PMG from *A. niger* showed optimum pH between 4 and 7 and highly esterified pectin (95%) was the best substrate (Koller and Neukom, 1967). The analysis of hydrolyzed products suggests *Aspergillus* produced only endo-PMG. Exo PMG has not been reported so far. A highly acidic pectinase (optimum pH 2.3) from *A. niger* has also been reported (Naidu and Panda, 1998b). The isolation of PMG and its biochemical characteristics need to be explored further for industrial applications.

2.2. Polygalacturonases

PG hydrolyzes the glycosidic linkages of polygalacturonates (pectates) by both exo and endo splitting mechanisms. Endo PGs act on the homogalacturonan backbone and break it into oligogalacturonates whereas exo-PGs break down polygalacturonates to di- and mono–galacturonates. PG activity can be determined by measuring the reducing sugars formed due to hydrolysis or by viscosity reduction method. However, the viscosity reduction method is less sensitive for exo PGs as the decrease in viscosity is relatively low. Cup plate method can also be used for estimating PG activity by viewing the clearing zones after staining with ruthenium red (Dingle *et al*, 1953; Truong *et al*, 2001). Endo PGs are widely distributed among fungi, bacteria and yeast. Endo PGs often occur in different forms having molecular weights in the range of 30–80 kDa and pI ranging between 3.8 and 7.6.

Most endo PGs have their optimum pH in the acidic range of 2.5–6.0 and an optimum temperature of 30 °C–50 °C (Singh and Rao, 2002; Takao *et al*., 2001). The K_m values of endo PGs are in the range of 0.14–2.7 mg/ml for pectate. PG shows no activity on highly methylated pectin. Exo PGs are widely distributed in *A. niger*, *Erwinia* sp. and in some plants such as carrots, peaches, citrus and apples (Pressey and Avants, 1975; Pathak and Sanwal, 1998). The molecular weight of exo PGs vary between 30–50 kDa and their pI ranges between 4.0 and 6.0.

Biochemical properties of pectinases in plants are crucial for food processing industries. The depolymerization of pectin by PG and other pectinases lead to a decrease in viscosity, which in turn, negatively affects the quality of tomato-based products. This can be prevented by selectively inactivating PG in the tomato by hydrostatic pressure, microwave heating or by ultrasound techniques. Recent studies of inactivation of PG by high pressure show promising results. PG I and PG II in tomatoes differ substantially in their thermal stability, PG II being more thermostable than PG I (Lopez *et al*., 1997). In another study, it has been reported that PG I is more thermostable than PG II (Anthon *et al*., 2002). The effect of temperature and pressure on the activity of purified tomato PG in the presence of pectins with various degrees of esterification was studied. The results showed a decrease in activity with an increase in pressure, at all temperatures. It has been reported that application of high pressure at ambient temperature caused approximately 70% decrease in PG activity. However, increasing the pressure from 300 to 700 MPa had no significant additional effect demonstrating the pressure resistance of PG (Krebbers *et al*., 2003). The residual PG activity was abolished at 90 °C and 700 MPa.

2.3. Pectin Lyases

Endo-PL degrades pectic substances in a random fashion yielding 4:5 unsaturated oligomethylgalacturonates and exo-PL has not been identified so far. Albersheim and coworkers first demonstrated transeliminative pectin depolymerization using pectin lyase from *A. niger* (Albersheim, 1966). Unsaturated oligogalacturonates can be estimated using spectrophotometeric method by measuring the increase in absorbance at 235 (molar extinction coefficient: $5.5 \times 10^{-5} M^{-1} cm^{-1}$) or using reducing sugar method or using thiobarbituric acid method (Nedjma *et al*., 2001). The measurement of viscosity reduction can also be used to measure the activity of PL but is predominantly used to determine whether the enzyme is endo or exo-splitting. Pectin lyases do not show absolute requirement of calcium for its activity except for *Fusarium* PL. However, it has been reported that PL activity can be stimulated in the presence of calcium. The molecular mass of PL lie in the range of 30 to 40 kDa (Soriano *et al*., 2005; Hayashi *et al*., 1997) except in the case of PL from *Aureobasidium pullulans* and *Pichia pinus* (~90 kDa). In general PL has been found to be active in acidic pH range of 4.0–7.0 although some reports show PL activity even in alkaline conditions (Soriano *et al*., 2005; Silva *et al.*, 2005). Isoelectric point has been found to be in the range of 3.5 for PL. The K_m values for PL are in the range between 0.1 mg/ml and 5 mg/ml respectively

depending on the substrate used (Sakiyama *et al*., 2001; Moharib *et al*., 2000). The thermal deactivation of PL from *A. niger* was modeled by first-order kinetics and found that the deactivation rate constant is minimum at pH 3.9 and 29 °C (Naidu and Panda, 2003). The effect of reaction and physical parameters on degradation of pectic substances was studied. The optimal amount of substrate and enzyme are 3.1 mg pectin and 1.67 U of PL, respectively, while the optimum pH and temperature are 4.8 and 35 °C, respectively (Naidu and Panda, 1999a). A substrate to enzyme ratio of 4 was the best for depolymerization of pectin by PL (Naidu and Panda, 1999b).

2.4. Polygalacturonate Lyase

Endo-PGL and exo-PGL are reported to degrade pectate by trans-elimination mechanism yielding 4,5 unsaturated oligogalacturonates, which can be quantified by methods described for PL. PGLs are found only in micro-organisms and they have an absolute requirement of calcium ions for activity. PGLs have an optimum pH near alkaline region (6–10), which is much higher than other pectinases (Singh *et al.*, 1999; Truong *et al*., 2001; Dixit *et al*., 2004). PGL are primarily produced by pathogenic bacteria belonging to *Erwinia*, *Bacillus* and certain other fungi like *Colletotrichum magna*, *Colletotrichum gloeosporiodes*, *Amylocota* sp. The molecular weight of PGL varies between 30–50 kDa except in the case of PGL from *Bacteroides* and *Pseudoalteromonas* (∼75 kDa) (McCarthy *et al*., 1985; Truong *et al.*, 2001). The optimum pH lies between 8.0 and 10.0 although PGL from *Erwinia* and *Bacillus licheniformis* were active even at pH 6.0 and 11.0 respectively. In general, the optimum temperature for PGL activity is between 30−40 °C. However, certain PGL from thermophiles have an optimum temperature between 50−75 °C. Pectates are good substrates for both endo and exo-PGL whereas chelating agents are inhibitors of PGL. Endo-PGL activity decreased with decrease in chain length of substrates and the rates are very slow when bi and trigalacturonates were substrates. However, exo-PGs do not show preference for size of substrate. In addition, there exists another class of enzymes called oligogalacturonate lyases (EC 4.2.2.6), which break down the oligogalacturonates and unsaturated oligogalacturonates by trans-elimination mechanism to remove unsaturated monomers from the reducing end of the substrate. These enzymes are predominantly produced by *Erwinia* and *Pseudomonas* sp. and the optimal pH is around 7.0.

2.5. Pectinesterase or Pectin Methylesterase

PE hydrolyzes the methoxy groups from 6-carboxyl group of galacturonan backbone of pectin. The product of degradation of pectin by PE is pectic acid, methanol and a proton from the ionization of newly formed carboxyl group. Pectin esterase activity can be determined in a pH-stat (Whitaker, 1984) or by titrating manually with a standard NaOH solution to maintain a constant pH or by observing the initial rate of decrease in pH from a fixed value (Nakagawa *et al*., 2000). Other ways

of determining the pectinesterase activity is by measuring the amount of methanol released by gas chromatography or by high performance liquid chromatography.

Pectinesterases are primarily produced in plants such as banana, citrus fruits and tomato and also by bacteria and fungi (Hasunuma *et al*, 2003). It has been reported that PE from fungi acts by a multi-chain reaction in which the methyl groups are removed in random fashion. However plant PE acts at non-reducing end or next to free carboxyl group and the methyl groups are removed by single-chain mechanism leading to the formation of blocks of deesterified galacturonate units (Froster, 1988). PE is more specific towards highly esterified pectic substances and shows no activity towards pectates. PE activity increases with increase in degree of esterification of substrate. The molecular weight of most microbial and plant PEs varies between 30–50 kDa (Hadj-Taieb *et al*, 2002; Christensen *et al*, 2002). The optimum pH for activity varies between 4.0 and 7.0 except for PE from *Erwinia* whose optimum pH is in alkaline region. Most PE has optimum temperature in the range 40−60 °C and a pI varying between 4.0 and 8.0. The values of K_m varies between 0.1–0.5 mg/ml. Industrially PE can be used to maintain the texture and firmness of processed fruit products and in clarification of fruit juices.

3. STRUCTURE AND FUNCTION OF PECTINASES

Three-dimensional structures of pectinases enable understanding of the molecular basis of enzyme mechanism, the role of individual amino acids in the active sites and also provide a rationale for structural differences between the enzymes that lead to very specific recognition of unique oligosaccharide sequences from a heterogeneous mixture in the plant cell wall. The information thus obtained allow for efforts to influence the functionality of these enzymes by rationally engineering novel or enhanced properties like faster processing and more specific cleavage patterns giving greater control of the structure of the processed pectins. Crystal structures of pectinases include members of all the major classes and the structure-function relationship studies of a few available complexes of pectinases with substrate/analogs could be considered as prototypes for related family members.

The first crystal structure of a pectinase was that of *Erwinia chrysanthemi* pectate lyase C (PelC) (Yoder *et al*, 1993). The same structural fold has subsequently been observed in other members of the pectinase family (Fig. 2). These include additional pectate lyases, *E. chrysanthemi* PelA (Thomas *et al*, 2002), PelE (Lietzke *et al*, 1994) and Pel9A (Jenkins *et al*, 2001); *Bacillus subtilis* Pel (Pickersgill *et al*, 1994) and high alkaline pectate lyase (Akita *et al*, 2000); two pectin lyases from *Aspergillus niger*, PLA (Mayans *et al*, 1997) and PLB (Vitali *et al.*, 1998); polygalacturonases, *Erwinia carotovora* polygalacturonase (Pickersgill *et al.*, 1998), *A. niger* endopolygalacturonase I and II (van Pouderoyen *et al*, 2003; van Santen *et al*, 1999) *Aspergillus aculeatus* polygalacturonase (Cho *et al*, 2001), *Stereum purpureum* endopolygalacturonase I (Shimizu *et al*, 2002) and *Fusarium*

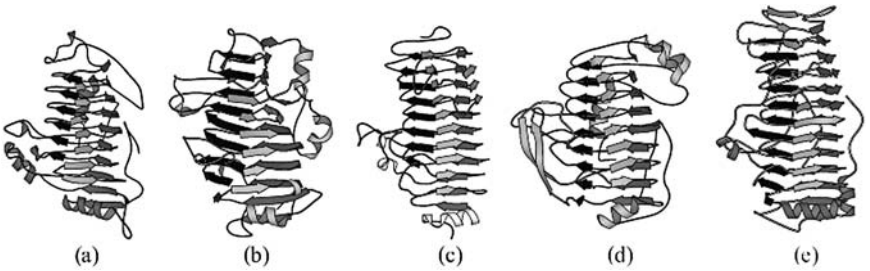


Figure 2. Examples of plant cell wall degradative enzymes that fold into a parallel β-helix motif. (a) *E. chrysanthemi* pectate lyase C. (b) Carrot pectin methylesterase (c) *F. moniliforme* endopolygalacturonase (d) *A. niger* pectin lyase A (e) *A. aculeatus* rhamnogalacturonase A. All structures are shown in an identical orientation with the N-terminal end at the bottom and the C-terminal end at the top. The three sets of β-sheets, PB1, PB2 and PB3 making up the fold are shaded differently, in increasing order of darkness of shade

moniliforme endopolygalacturonase (Federici *et al*., 2001); pectin methylesterases from *Daucus carota* (Johansson *et al*., 2002), *Lycopersicon esculentum* (Di Matteo *et al.*, 2005) and PemA from *E. chrysanthemi* (Jenkins *et al*., 2001) and *Aspergillus aculeatus* rhamnogalaturonase A (Peterson *et al*., 1997).

Each structure consists of a single domain of parallel β-strands folded into a large right-handed cylinder. The domain fold, termed the parallel β-helix, is compatible with all accepted structural rules, albeit in a unique manner. The central cylinder consists of seven to nine complete helical turns and is prism shaped due to the unique arrangement of three parallel β-strands in each turn of the helix. The strands of consecutive turns line up to form three parallel β-sheets called PB1, PB2 and PB3. PB1 and PB2 form an antiparallel β sandwich, while PB3 lies approximately perpendicular to PB2. Although the mechanism of pectic cleavage differs for the esterases, hydrolases and lyases, the substrate binding sites as deduced from structures, sequence similarity and site directed mutagenesis studies (Kita *et al.*, 1996), are all found in a similar location within a cleft formed on the exterior of the parallel β-helix between one side of PB1 and the protruding loops (Fig. 3). The structural differences in the loops are believed to be related to subtle differences in the enzymatic and maceration properties.

The next sections will discuss the structure-function relationships identified from structural studies of four industrially important pectinases, namely, PG, PGL, PL and PE. There is presently no structure of a representative from the PMG family.

3.1. Polygalacturonases

The endoPGs are inverting glycosidases that invert the anomeric configuration of the products during the reaction. In this mechanism, the hydrolysis proceeds by a general acid catalyst donating a proton to the glycosidic oxygen and a catalytic base guiding the nucleophilic attack of a water molecule on the anomeric carbon

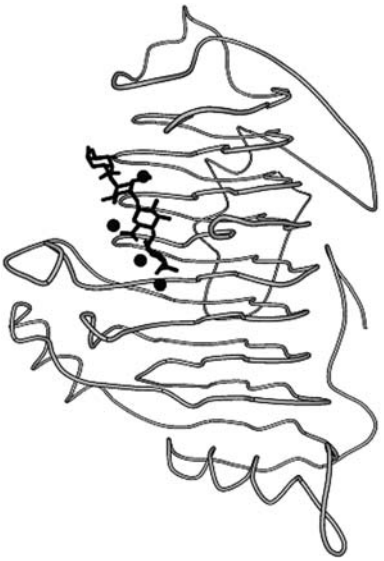


Figure 3. E. chrysanthemi pectate lyase C in complex with a cell wall fragment. The ordered tetraGal*p*A fragment is shown in stick representation while the four Ca2+ ions are shown as spheres. The entire protein backbone is shown. The substrate binding site is made up a large cleft formed by one of the parallel β-sheet, PB1, and the loops of both sides of PB1. (See Fig. 2 legend)

of the galacturonate moiety bound at the −1 subsite. The crystal structures of native *S. purpureum* endo PG I and that of the ternary product complexes with two molecules of galacturonate provided experimental evidence of the substrate binding mechanism, active site architecture as well as the reaction mechanism (Shimizu *et al*., 2002). The interactions with bound uronates identified as β-D-galactopyranuronic acid (Gal*p*A) and β-D-galactofuranuronic acid (Gal*f*A) of the ternary complex are shown in Fig. 4a and 4b, respectively. The Gal*p*A binding site is believed to be the +1 subsite, because its location is at the reducing end side of a proposed catalytic residue Asp173 whereas the binding site for Gal*f*A is at the proposed −1 subsite, because it is located on the opposite side of the +1 subsite. In site-directed mutagenesis studies of *A. niger* endoPG II, the replacement of charged residues His195, Arg226, Lys228, and Tyr262 led to 10-fold or greater increases in the K_m value (Armand *et al*., 2000; Pages *et al*., 2000) thus confirming the importance of the carboxy group recognition in subsite +1 for productive substrate binding. In contrast, the replacement of Asp173 caused only a 2-fold increase in the K_m value, but greatly decreased the K_{cat} value (Armand *et al*., 2000). Asp173 is expected to serve as a general acid catalyst that donates a proton to the glycosidic oxygen. Thus, the tight binding of the substrate to subsite +1 is due to the electrostatic interactions between the carboxy group and the basic residues and the precise recognition of the galactose epimer. In the case of Gal*f*A binding in the ternary complex, the carboxy group is recognized by a conserved structural

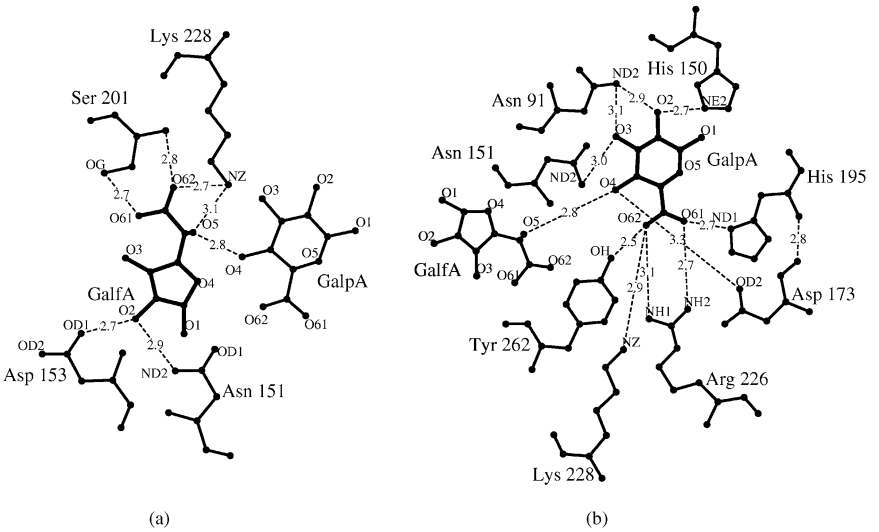


Figure 4. (a) and (b). Schematic representation of the interactions between *Stereum purpureum* endoPG I and two galacturonates in the ternary complex. The dotted lines show hydrogen bonds and electrostatic interactions, and their distances are indicated in angstroms

motif. In both the −1 and +1 subsites, binding of the carboxy group is considered an important mechanism of substrate recognition. This probably accounts for the fact that endo PGs are able to cleave only free polygalacturonate and not the methylesterified substrate (Shimizu *et al*., 2002).

3.2. Polygalacturonate Lyase

All proteins in the PGL family are believed to share a similar enzymatic mechanism. The enzyme randomly cleaves pectates by a β-elimination mechanism, generating primarily a trimer end-product with a 4,5-unsaturated bond in the galacturonosyl residue at the non-reducing end (Preston *et al*., 1992). The β-elimination reaction in pectolytic cleavage involves three steps: neutralization of the carboxyl group adjacent to the scissile glycosidic bond, abstraction of the C5 proton, and transfer of the proton to the glycosidic oxygen. Among the structures of members of the PGL superfamily, the Michaelis complex of a catalytically inactive R218K PGL mutant, PelC from *E. chrysanthemi* and a plant cell wall fragment (a penta Gal*p*A substrate) reveals important details regarding the enzymatic mechanism (Scavetta *et al.*, 1999).

Structural studies of PelC at various pH and Ca^{2+} concentrations have shown that Ca^{2+} binding is essential to the *in vitro* activity of PelC and that the Ca^{2+} ion has multiple functions (Pickersgill *et al*., 1994; Herron *et al*., 2003). As shown in Fig. 3, four well-ordered Gal*p*A units interact with PelC in a groove where

the reducing end Gal*p*A is located at the protein–solvent border, and the non-reducing end, Gal*p*A, lies near a Ca^{2+}. Electrostatic interactions dominate, with the negatively charged uronic acid moieties primarily interacting with the four Ca^{2+} ions found in the R218K complex with penta Gal*p*A. The Ca^{2+} ions link not only the oligosaccharide to the protein but also adjacent uronic acid moieties within a single pectate strand. The observed Ca^{2+} positions are very different from the interstrand Ca^{2+} ions postulated to link PGA helices together (Walkinshaw and Arnott, 1981a,b). The carboxyl oxygens of Gal*p*A2, Gal*p*A3 and Gal*p*A4 interact with Arg245, Lys190 and Lys172 respectively. Lys172 and Lys190 are highly conserved in PGL, but PL that binds a neutral methylated form of pectate, lack both amino acids thus providing a probable structural basis for differences in substrate specificities between PL and PGL.

Since PelC and subfamily members are reported to yield a trimer as the primary unsaturated end-product and all interactions of Gal*p*A3 and Gal*p*A4 with PelC involve highly conserved and invariant amino acids within the PGL family, it has been postulated that the scissile bond occurs between Gal*p*A3 and Gal*p*A4 (Scavetta *et al.*, 1999). Although it is highly unusual for an arginine to act as a general base during catalysis, Arg218 has been proposed to be the group responsible for proton abstraction and transfer based on its orientation and interactions in native PelC and other data including impairment of catalysis in the R218K mutant and sequence conservation among the PGL superfamily (Scavetta *et al.*, 1999). A review on the structure and function of PelC discusses the pathogenesis mechanism at the molecular level (Herron *et al.*, 2000).

Structural studies of PGLs have also revealed that they do not always adopt the characteristic parallel β-helix fold. The structures of a PGL (PelA) from *Azospirillum irakense* (Novoa De Armas *et al.*, 2004) and that of the catalytic module of the *Cellvibrio japonicus* PGL (Pel10Acm) (Brown *et al.*, 2001) adopt a predominantly α-helical structure with irregular coils and short β-strands. They show two 'domains' with the interface between them being a wide-open central groove in which the active site is located. Both belong to different families of polysaccharide lyases in the carbohydrate-active enzymes (CAZY) classification (Coutinho and Henrissat, 1999, http://afmb.cnrs-mrs.fr/CAZY/). However, comparison of the structures of Pel10Acm GalA3/Ca^{2+} complex with the *E. chrysanthemi* inactive mutant R218K PelC complex with GalA4/Ca^{2+}, reveals an essentially identical disposition of six active site groups despite no topological similarity between these enzymes. Identification of common coordination of the –1 and +1 subsite saccharide carboxylate groups by a protein-liganded Ca^{2+} ion, the positioning of an arginine catalytic base in close proximity to the α-carbon hydrogen, numerous other conserved enzyme–substrate interactions and mutagenesis data suggest a common polysaccharide *anti*-β-elimination mechanism for both families (Brown *et al.*, 2001). The absence of sequence homology between distinct families of pectate lyases suggests that such catalytically similar enzymes have evolved independently and may reflect their different functions in nature.

3.3. Pectin Lyase

Enzymatic cleavage in PL occurs via the same β-elimination reaction as seen for PGLs. However, in contrast to PGL, PL are specific for highly methylated forms of the substrate and do not require Ca^{2+} for activity. Crystal structures available for the apo- forms of pectin lyase A (Mayans *et al*., 1997) and pectin lyase B (Vitali *et al*., 1998) from *A. niger* show that both adopt the parallel β-helix fold (Fig. 2d) and are structurally almost identical. Comparison of structures of PGL and PL show that although they share many structural features, there appears to be remarkable divergence in the substrate binding clefts and catalytic machinery reflecting differences in their substrate specificities.

The divergence in substrate specificity comes from two factors. Firstly, the loops making the active side cleft in PL is much longer and of a more complex conformation that encompasses two β-strands forming an antiparallel β-sheet (Figs. 2a and 2d). Secondly, the putative active site of PL exhibits a cleft that is predominantly aromatic comprising of four tryptophans and three tyrosines that contribute to the architecture of the active site. In contrast, as discussed earlier, PGL presents a binding cleft rich in charged amino acids with no aromatic residues and uses Ca^{2+} in catalysis as an aid for the activation of the C5 proton. A conserved Arg176 in PL is proposed to play a similar catalytic role like that of Ca^{2+} in PGL. The comparison of PL and PGL structures also reveal that in PL, regions of strong negative electrostatic potential envelops the aromatic substrate binding cleft thus contributing to repulsion of negative charged pectate which is not a substrate whereas in PGL, the electrostatic field around the substrate binding site is a ribbon of positive potential that attracts the demethylated pectate substrate. Thus, substrate specificity is a consequence of hydrophobicity of the binding cleft and long-range electrostatic effects. In conclusion, although these enzymes share a common fold and related mechanisms, their strategies for substrate recognition and binding are completely different (Mayans *et al*., 1997).

3.4. Pectin Methylesterase

PEs catalyze pectin deesterification by hydrolysis of the ester bond of methylated α-(1-4)-linked D-galacturonosyl units, producing a negatively charged polymer and methanol. Among the PE structures are those of *E. chrysanthemi* PemA and a PE from carrot. They both have similar structures and belong to the family of parallel β-helix proteins with major differences in loops making up the substrate binding cleft (Fig. 2b). The putative active site of PE was deduced from mapping of sequence conservation among PEs onto the structure. The substrate binding cleft is located in a location similar to that of the active site and substrate binding cleft of PGL and PL. The central part of the cleft is lined by several conserved aromatic amino acids, especially on the exposed side of PB1. The active site of PE is located in the long shallow cleft lined by two absolutely conserved aspartic acid residues in the center, Asp136 and Asp157 in carrot PE. The following mechanism of action

has been suggested for carrot PE. Asp157 acts as the nucleophile for the primary attack on the carboxymethyl carbonyl carbon while Asp136 may act as an acid in the first cleavage step, where methanol is leaving. Asp136 could then act as a base, in the next step extracting a hydrogen from an incoming water molecule to cleave the covalent bond between the substrate and Asp157 to restore the active site.

4. APPLICATIONS OF PECTINASES

Industrial applications of pectinases have been reviewed by different authors (Kashyap *et al*., 2001; Hoondal *et al*., 2002; Naidu and Panda, 1998a; Gummadi and Panda, 2003; Gummadi and Kumar, 2005). Pectinases from microbial and plant sources have thus received a lot of attention. From an industrial point of view, pectinases are classified into two types, acidic and alkaline pectinases. The acidic pectinases have extensive applications in extraction and clarification of both

Table 2. Industrial applications of pectinases

Application	Purpose	Reference
Cloud stabilization	To precipitate hydrocolloid matter present in fruit juices	Rebeck, 1990; Grassin and Fauquembergue, 1996
Fruit juice clarification	Degradation of cloud forming pectic substances. Hence, the juice can be easily filtered and processed	Rombouts and Pilnik, 1986; Alkorta *et al.,* 1998
Extraction of juice and oil	To overcome the difficulty in pressing fruit pulp to yield juice and oil	Kilara, 1982; Pilnik and Voragen, 1993
Maceration	To break down the vegetable and fruit tissues to yield pulpy products used as base material for juices, nectar as in the case of baby foods, pudding and yogurt	Fogarty and Kelly, 1983
Liquefaction	To break down fermentable plant carbohydrates to simple sugars by enzymes	Beldman *et al.,* 1984
Gelation	To use in gelling low-sugar fruit products	Spiers *et al.,* 1985
Wood preservation	To prevent the wood from infection by increasing the permeability of wood preservative	Fogarty and Ward, 1973
Retting of fiber crops	To release fiber from the crops by fermenting with micro-organisms, which degrade pectin	Henriksson *et al.,* 1999
Degumming of fiber crops	To remove the ramie gum of ramie fiber	Gurucharanam and Deshpande, 1986; Zheng *et al.,* 2001
Waste water treatment	To degrade pectic substances in waster water from citrus processing industries	Peterson, 2001; Tanabe *et al.,* 1987
Coffee and tea fermentation	To remove the mucilage coat in coffee bean. To enhance the tea fermentation and foam forming property of tea	Carr, 1985; Godfrey, 1985

sparkling clear juice (apple, pear, grapes and wine) and cloudy (lemon, orange, pineapple and mango) juices and maceration of plant tissues (Table 2). Additionally, acidic pectinases are useful in the isolation of protoplasts (Takebe *et al*, 1968) and saccharification of biomass (Beldman *et al*, 1984). Alkaline pectinases have potential applications in cotton scouring, degumming of plant fibers to improve the quality of fiber, coffee and tea fermentation, paper industry and for purification of plant viruses (Salazar and Jayasinghe, 1999).

ACKNOWLEDGEMENTS

Gummadi acknowledges financial support from Department of Science and Technology, India.

REFERENCES

Akita, M., Suzuki, A., Kobayashi., T., Ito, S. and Yamane, T. (2000) Crystallization and preliminary X-Ray analysis of High-Alkaline Pectate Lyase. Acta Crystallogr. D. *56*, 749–750.

Albersheim, P. (1966) Pectin lyase from fungi. Methods enzymol. *8*, 628–631.

Alkorta, I., Garbisu, G., Llama, M.J. and Serra, J.L. (1998) Industrial applications of pectic enzymes: A review. Proc. Biochem. *33*, 21–28.

Anthon, G.E., Sekine, Y., Watanabe, N. and Barrett. D.M. (2002) Thermal Inactivation of Pectin Methylesterase, Polygalacturonase, and Peroxidase in Tomato Juice. Agric. Food Chem. *50*, 6153–6159.

Armand, S., Wagemaker, M.J., Sanchez-Torres, P., Kester, H.C., van Santen, Y., Dijkstra, B.W., Visser, J. and Benen, J.A. (2000) The active site topology of *Aspergillus niger* endopolygalacturonase II as studied by site-directed mutagenesis. J. Biol. Chem. *275*, 691–696.

Beldman, G., Rombouts, F.M., Voragen, A.G.J. and Pilnik, W. (1984) Application of cellulase and pectinase from fungal origin for the liquifaction and sachharification of biomass. Enzyme Microbiol. Technol. *6*, 503–507.

Brown, I.E., Mallen, M.H., Charnock, S.J., Davies, G.J. and Black, G.W. (2001) Pectate lyase 10A from *Pseudomonas cellulosa* is a modular enzyme containing a family 2a carbohydrate-binding module. Biochem. J. *355*, 155–165.

Carr, J.G. (1985) Tea, coffee and cocoa. In: Wood, B.J.B. (Eds.), Microbiology of Fermented Foods, vol. II. Elsevier Applied Science, London, pp. 133–154.

Chin, L., Ali, Z.M. and Lazan, H. (1999) Cell wall modifications, degrading enzymes and softening of carambola fruit during ripening. J. Experimental Botany. *50*, 767–775.

Cho, S.W, Lee, S and Shin, W. (2001) The X-ray structure of *Aspergillus aculeatus* polygalacturonase and a modeled structure of the polygalacturonase-octagalacturonate complex. J. Mol. Biol. *311*, 863–878.

Christensen. T.M., Nielsen, J.E., Kreiberg, J.D., Rasmussen, P. and Mikkelsen, J.D. (2002) Pectin methyl esterase from orange fruit: characterization and localization by *in-situ* hybridization and immunohistochemistry. Planta. *206*, 493–503.

Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-active enzymes: an integrated database approach. In: Gilbert, H.J., Davies, G., Henrissat, B. and Svensson B. (Eds.), Recent Advances in Carbohydrate Bioengineering, The Royal Society of Chemistry, Cambridge, pp. 3–12.

Di Matteo, A., Giovane, A., Raiola, A., Camardella, L., Bonivento, D., De Lorenzo, G., Cervone, F., Bellincampi, D. and Tsernoglou, D. (2005) Structural basis for the interaction between pectin methylesterase and a specific inhibitor protein. Plant Cell.*17*, 849–858.

Dingle, J., Reid, W.W. and Solomons, G.L. (1953) The enzymatic degradation of pectin and other polysaccharides. II. Application of the 'Cup-plate' assay to the estimation of enzymes. J. Sci. Food Agric. *4*, 149–155.

DiPietro, A. and Roncero, M.I. (1996) Purification and characterization of an exo-polygalacturonase from the tomato vascular wilt pathogen *Fusarium oxysporum* f.sp. *lycopersici.* FEMS Microbiol. Lett. *145*, 295–299.

Dixit, V.S., Kumar, A.R., Pant, A. and Khan, M.I. (2004) Low molecular mass pectate lyase from *Fusarium moniliforme*: similar modes of chemical and thermal denaturation. Biochem. Biophys. Res. Commun. *315*, 477–484.

Federici, L., Caprari, C., Mattei, B., Savino, C., Di Matteo, A., De Lorenzo, G., Cervone, F. and Tsernoglou, D. (2001) Structural requirements of endopolygalacturonase for the interaction with PGIP (polygalacturonase-inhibiting protein). Proc. Natl. Acad. Sci. *98*, 13425–13430.

Fogarty, W.M. and Ward, O.P. (1973) Pectinases and pectic polysaccharides. Prog. Ind. Microbiol. *13*, 61–119.

Fogarty, W.M. and Kelly, C.T. (1983) Pectic enzymes. In: Fogarty, M.W. (Eds.), Microbial Enzymes and Biotechnology. Applied Science Publishers, London, pp. 131–182.

Forster, H. (1988) Pectinesterase from *Phytophthora infestans*. Methods Enzymol. *161*, 355–357.

Godfrey, A. (1985) Production of industrial enzymes and some applications in fermented foods. In: Wood BJB (Eds.) Microbiology of fermented foods, Vol 1. Elsevier, London, pp 345-371.

Grassin, C. and Fauquembergue, P. (1996) Fruit juices. In: Godfrey, T., West, S. (Eds.), Industrial Enzymology, second ed. Stockholm Press, New York, pp. 225-264.

Gummadi, S.N. and Kumar, D.S. (2005) Microbial pectic transeliminases. Biotechnol. Lett. *27*, 451–458.

Gummadi, S.N. and Panda, T. (2003) Purification and biochemical properties of microbial pectinases: a review. Proc. Biochem. *38*, 987–996.

Gurucharanam, K. and Deshpande, K.S. (1986) Polysaccharides of *Curbularia lunata*: Use in degumming of ramie fibers. Ind. J. Phytopathol. *39*, 385–389.

Hadj-Taieb, N., Ayadi, M., Trigui, S., Bouabdollah, F. and Gargouri, A. (2002) Hyper production of pectinase activities by fully constitutive mutant (CT 1) of *Penicillium occitanis*. Enzyme Microbial. Technol. *30*, 662–666.

Hasunuma, T., Fukusaki, E.I. and Kobayashi, A. (2003) Methanol production is enhanced by expression of an *Aspergillus niger* pectin methylesterase in tobacco cells. J. Biotechnol. *106*, 45–52.

Hayashi, K., Inoue, Y., Shiga, M., Sato, S., Takano, R., Hirayae, K., Hibi, T. and Hara, S. (1997) Pectinolytic enzymes from *Pseudomonas marginalis* MAFF 03-01173. Phytochemistry.*45*, 1359–1363.

Henriksson, G., Akin, D.E., Slomczynski, D. and Eriksson, K.E.L. (1999) Production of highly efficient enzymes for flax retting by *Rhizomucor pusillus*. J. Biotechnol. *68*, 115–123.

Herron, S.R., Benen, J.A., Scavetta, R.D., Visser, J. and Jurnak, F. (2000) Structure and function of pectic enzymes: virulence factors of plant pathogens. Proc. Natl. Acad. Sci. U S A. *97*, 8762–8769.

Herron, S.R., Scavetta, R.D., Garrett, M., Legner, M. and Jurnak, F. (2003) Characterization and implications of Ca^{2+} binding to pectate lyase C.J. Biol. Chem. *278*, 12271–12277.

Hoondal, G.S., Tewari, R.P., Tewari, R., Dahiya, N. and Beg, Q.K. (2002) Microbial alkaline pectinases and their industrial applications a review. Appl. Microbiol. Biotechnol. *59*, 409–418.

Jenkins, J., Mayans, O., Smith, D., Worboys, K. and Pickersgill, R.W. (2001) Three-dimensional structure of *Erwinia chrysanthemi* pectin methylesterase reveals a novel esterase active site. J. Mol. Biol. *305*, 951–60.

Johansson, K., El-Ahmad, M., Friemann, R., Jornvall, H., Markovic, O. and Eklund, H. (2002) Crystal structure of plant pectin methylesterase. FEBS Lett.*51*, 243–249.

Kashyap, D.R., Vohra, P.K., Chopra, S., Tewari, R. (2001) Applications of pectinases in commercial sector: a review. Biores. Technol. *77*, 215–227.

Kilara, A. (1982) Enzymes and their uses in the processed apple industry: A Review. Proc. Biochem. *23*, 35–41.

Kita, N., Boyd, C.M., Garrett, M.R., Jurnak, F. and Keen, N.T. (1996) Differential effect of site-directed mutations in PelC of pectate lyase activity, plant tissue maceration, and elicitor activity. J. Biol. Chem. *271*, 26529–26535.

Koller, A. and Neukom, H. (1967) Onteruschimgen uber den pektolytischen enzyme von *Aspergillus niger*. Mitt Geb Debensmittelunter Hug. *58*, 512.

Krebbers, B., Matser, A.M., Hoogerwerf, S.W., Moezelaar, R., Tomassen, M.M., Robert, W. and Vanden Berg, R.W. (2003) Combined high-pressure and thermal treatments for processing of tomato puree: evaluation of microbial inactivation and quality parameters. Innovative Food Sci. Emerg. Technol. *4*, 377–385.

Lang, C. and Dörenberg, H. (2000) Perspective in the biological function and the technological applications of polygalacturonases. Appl. Microbiol. Biotechnol. *53*, 366–375.

Lietzke, S.E., Keen, N.T., Yoder, M.D. and Jurnak, F. (1994) The Three-Dimensional Structure of Pectate Lyase E, a Plant Virulence Factor from *Erwinia chrysanthemi*. Plant Physiol.106, 849–862.

Lopez, P., Sanchez, A.C., Vercet, A. and Burgos, J. (1997) Thermal resistance of tomato polygalacturonase and pectinmethylesterase at physiological pH. Z. Lebensm.-Unters.-Forsch. A. *204*, 146–150

Mayans, O., Scott, M., Connerton, I., Gravesen, T., Benen, J., Visser, J., Pickersgill, R. and Jenkins, J. (1997) Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate binding clefts of pectin and pectate lyases. Structure. *5*, 677–689.

Mayans, O., Scott, M., Connerton, I., Gravesen, T., Benen, J., Visser, J., Pickersgill, R. and Jenkins, J. (1997) Two crystal structures of pectin lyase A from *Aspergillus* reveal a pH driven conformational change and striking divergence in the substrate binding clefts of pectin and pectate lyases. Structure. *5*, 677–689.

McCarthy, R.E., Kotarski, S.F. and Salyers, A.A. (1985) Location and characteristics of enzymes involved in the breakdown of polygalacturonic acid by *Bacteroides thetaiotaomicron.* J. Bacteriol. *161*, 493–499.

Moharib, S.A., El-Sayed, S.T. and Jwanny, E.W. (2000) Evaluation of enzymes produced from yeast. Nahrung. *44*, 47–51.

Naidu, G.S.N. and Panda, T. (1998a) Production of pectolytic enzymes – a review. Bioproc. Eng. *19*, 355–361.

Naidu, G.S.N. and Panda, T. (1998b) Application of response surface methodology to evaluate some aspects of pectolytic enzymes from *Aspergillus niger*. Biochem. Engg. J. *2*, 71–77.

Naidu, G.S.N. and Panda, T. (1999a) Performance of pectolytic enzymes during hydrolysis of pectic substances under assay conditions: a statistical approach. Enzyme Microbial. Technol. *25*, 116–124.

Naidu, G.S.N. and Panda, T. (1999b) Behaviour of pectinases under process conditions. Bioproc. Eng. *21*, 397–400.

Naidu, G.S.N. and Panda, T. (2003) Studies on pH and thermal deactivation of pectolytic enzymes from *Aspergillus niger*. Biochem. Eng. J. *16*, 57–67.

Nakagawa, T., Miyaji, T., Yurimoto, H., Sakai, Y., Kato, N. and Tomizuka, M.A. (2000) Methylotrophic pathway participates in pectin utilization by *Candida boidinii*. Appl. Environ. Microbiol. *66*, 4253–4257.

Nedjma, M., Hoffmann, N. and Belari, A. (2001) Selective and sensitive detection of pectin lyase activity using a colorimetric test: Application to the screening of micro-organisms possessing pectin lyase activity. Anal. Biochem. *291*, 290–296.

Novoa De Armas, H., Verboven, C., De Ranter, C., Desair, J., Vande Broek, A., Vanderleyden, J. and Rabijns, A. (2004) *Azospirillum irakense* pectate lyase displays a toroidal fold. Acta Crystallogr D. *60*, 999–1007.

Pages, S., Heijne, W.H., Kester, H.C., Visser, J. and Benen, J.A. (2000) Subsite mapping of *Aspergillus niger* endopolygalacturonase II by site-directed mutagenesis. J. Biol. Chem. *275*, 29348–29353.

Pathak, N. and Sanwal, G.G. (1998) Multiple forms of polygalacturonase from banana fruits. Phytochem. *48*, 249–255.

Petersen, S. (2001) Enzymes to upgrade plant nutrients. Feed Mix. *9*, 12–15.

Petersen, T.N., Kauppinen, S. and Larsen, S. (1997) The crystal structure of a rhamnogalacturonase A from *Aspergillus aculeatus*: a right-handed parallel beta helix. Structure. *5*, 533–544.

Pickersgill, R., Jenkins, J., Harris, G., Nasser, W. and Robert-Baudouy, J. (1994) The structure of *Bacillus subtilis* pectate lyase in complex with calcium. Nat. Struct. Biol.1, 717–723.

Pickersgill, R., Smith, D., Worboys, K. and Jenkins, J. (1998) Crystal structure of polygalacturonase from *Erwinia carotovora ssp. carotovora*. J. Biol. Chem .273, 24660–24664

Pilnik, W. and Voragen, A.G.J. (1993) Pectic enzymes in fruit juice and vegetable juice manufacture. In: Reeds, G. (Eds.), Food and Science Technology, Enzymes in Food Processing. Academic Press, New York, pp. 363–399 .

Pressey, R. and Avants, J.K. (1975) Modes of action of carrot and peach exopolygalacturonases. Phytochem. *14*, 957–961.

Preston, J.F., Rice, J.D., Ingram, L.O. and Keen, N.T. (1992) Differential depolymerization mechanisms of pectate lyases secreted by *Erwinia chrysanthemi* EC16. J. Bacteriol. *174*, 2039–2042.

Rebeck, H. (1990) Processing of citrus juices. In: Hick, D. (Eds.), Production and Packaging of Non-Carbohydrate Fruit Juices and Fruit Beverages. Van Nosrand Reinhold, New York.

Rombouts, F.M. and Pilnik, W. (1986) Pectinases and other cell-wall degrading enzymes of industrial importance. Symbiosis *2*, 79–89.

Sakiyama, C.C.H., Paula, E.M., Pereira, P.C., Borges, A.C. and Silva, D.O. (2001) Characterization of Pectin lyase produced by an endophytic strain isolated from coffee cherries. Lett. Appl. Microbiol. 33, 117–121.

Salazar, L. and Jayasinghe, U. (1999) Fundamentals of purification of plant viruses. In: Techniques in plant virology CIP Training manual 5.0, Virus purification, International potato society, Peru, pp. 1–10. ISSN 0256 6311.

Scavetta, R.D., Herron, S.R., Hotchkiss, A.T., Kita, N., Keen, N.T., Benen, J.A., Kester, H.C., Visser, J. and Jurnak, F. (1999) Structure of a plant cell wall fragment complexed to pectate lyase C. Plant Cell. *11*, 1081–1092.

Shimizu, T., Nakatsu, T., Miyairi, K., Okuno, T. and Kato, H. (2002) Active-site architecture of endopolygalacturonase I from *Stereum purpureum* revealed by crystal structures in native and ligand-bound forms at atomic resolution. Biochemistry. *41*, 6651–6659.

Silva, D., Tokuioshi, K., Silva, E.D., SilvaMartins, R.D. and Gomes, E. (2005) Production of pectinases by solid state fermentation with *Penicillium viridicatum* RFC3. Proc. Biochem. *40*, 2885–2889.

Singh, S.A. and Rao, A.G.A. (2002) A simple fractionation protocol for, and a comprehensive study of the molecular properties of two major endopolygalacturonases from *Aspergillus niger*. Biotechnol. Appl. Biochem. *35*, 115–123.

Singh, S.A., Plattnera, H. and Diekmann, H. (1999) Exopolygalacturonate lyase from a thermophulic *Bacillus* sp. Enzyme Microbiol. Technol. *25*, 420–425.

Soriano, M., Diaz, P. and Pastor, F.I.J. (2005) Pectinolytic systems of two aerobic sporogenous bacterial strains with high activity on pectin. Curr. Microbio. *50*, 114–118.

Sozzi-Quiroga, G.O. and Fraschina, A.A. (1997) Evaluation of sensory attributes and biochemical parameters in transgenic tomato fruit with reduced polygalacturonase activity. Food Sci. Technol. Int. *3*, 93–102.

Spiers, J.M., Braswell, J.M. and Hegwood, C.P. (1985) Establishment of rabbiteye blueberries. Mississippi Agr. Expt. Sta. Bul. 941.

Takao, M., Nakaniwa, T., Yoshikawa, K., Terashita, T. and Sakai, T. (2001) Molecular cloning, DNA sequence, and expression of the gene encoding for thermostable pectate lyase of thermophilic *Bacillus* sp. TS 47. Biosci. Biotechnol. Biochem. *65*, 322–329.

Takebe, I., Otsuki, Y., Aoki, S. (1968) Isolation of tobacco mesophyll cells in intact and active state. Plant Cell Physiol. *9*, 115.

Tanabe, H., Yoshihara, K., Tamura, K., Kobayashi, Y., Akamatsu, I., Niyomwan, N. and Footrakul, P. (1987) Pretreatment of pectic wastewater from orange canning process by an alkalophilic *Bacillus* sp. J. Fermentation Technol. *65*, 243–246.

Ten Have, A., Mulder, W., Visser, J. and Van Kan, J.A.L. (1998) The endopolygalacturonase gene Bcpg1 is required for full virulence of *Botrytis cinerea*. Mol. Plant Microbe. Interact. *11*, 1084–1097.

Thomas, L.M., Doan, C.N., Oliver, R.L. and Yoder, M.D. (2002) Structure of pectate lyase A: comparison to other isoforms. Acta Crystallogr D. *58*, 1008–1015.

Truong, L.V., Tuyen, H., Helmke, E., Binh, L.T. and Schweder, T. (2001) Cloning of two pectate lyase genes from the marine Antarctic bacterium *Pseudoalteromonas haloplanktis* strain ANT/505 and characterization of the enzymes. Extremophiles. *5*, 35–44.

Van Pouderoyen, G., Snijder, H.J., Benen, J.A. and Dijkstra, B.W. (2003) Structural insights into the processivity of endopolygalacturonase I from *Aspergillus niger*. FEBS Lett. *554*, 462–466.

Van Santen, Y., Benen, J.A.E., Schroter, K.H., Kalk, K.H., Armand, S., Visser, J. and Dijkstra, B.W. (1999) 1.68-A crystal structure of endopolygalacturonase II from *Aspergillus niger* and identification of active site residues by site-directed mutagenesis J. Biol. Chem. *274*, 30474–30480.

Vincent-Sealy, L.V., Thomas, J.D., Commander, P. and Salmond, G.P. (1999) *Erwinia carotovora* DsbA mutants: evidence for a periplasmic-stress signal transduction system affecting transcription of genes encoding secreted proteins. Microbiology. *145*, 1945–1958.

Vitali, J., Schick, B., Kester, H.C.M., Visser, J. and Jurnak, F. (1998) The three dimensional structure of *Aspergillus niger* pectin lyase B at 1.7-A resolution. Plant Physiol. *116*, 69–80.

Walkinshaw, M.D. and Arnott, S. (1981a) Conformations and interactions of pectins. I. X-ray diffraction analysis of sodium pectate in neutral acidified forms. J. Mol. Biol. *153*, 1055–1073.

Walkinshaw, M.D. and Arnott, S. (1981b) Conformations and interactions of pectins. II. Models of junction zones in pectinic acid and calcium pectate gels. J. Mol. Biol. *153*, 1075–1085.

Whitaker, J.R. (1984) Pectic substances, pectic enzymes and haze formation in fruit juices. Enzyme Microbial. Technol. *6*, 341–347.

Whitaker, J.R. (1991) Microbial Enzymes and Biotechnology, Second edition. In: Fogarty, M. Kelly, C.T. (Eds.), Elsevier Applied Science, London, pp. 133–176.

Yoder, M.D., Keen, N.T. and Jurnak, F. (1993) New domain motif: structure of pectate lyase C, a secreted plant virulence factor. Science. *260*, 1503–1507.

Zheng, L., Du, Y. and Zhang, J. (2001) Degumming of ramie fibers by alkalophilic bacteria and their polysaccharide degrading enzymes. Bioresour. Technol. *78*, 89–94.

CHAPTER 8

α-L-RHAMNOSIDASES: OLD AND NEW INSIGHTS

PALOMA MANZANARES[1*], SALVADOR VALLÉS[1], DANIEL RAMÓN[1,2] AND MARGARITA OREJAS[1]

[1]*Departamento de Biotecnología de Alimentos, Instituto de Agroquímica y Tecnología de Alimentos, CSIC, Paterna, Valencia, Spain and*

[2]*Departamento de Medicina Preventiva y Salud Pública, Ciencias de los Alimentos, Toxicología y Medicina Legal, Universitat de València, Valencia, Spain*

[*]*pmanz@iata.csic.es*

1. INTRODUCTION

L-Rhamnose is a component of plant cell wall pectic polysaccharides (Mutter *et al.*, 1994; Ridley *et al*, 2001), glycoproteins (Haruko and Haruko, 1999) and secondary metabolites such as anthocyanins (Renault *et al*, 1997), flavonoids (Bar-Peled *et al*, 1991) and triterpenoids (Friedman and McDonald, 1997). It has also been found in bacterial heteropolysaccharides (Hashimoto and Murata, 1998), rhamnolipids (Ochsner *et al*, 1994) and in the repeating units of the O-antigen structure of the lipopolysaccharide component of bacterial outer membranes (Chua *et al.*, 1999). Some rhamnosides are important bioactive compounds, *e.g.* cytotoxic saponins (Bader *et al*, 1998; Yu *et al*, 2002), antifungal plant glycoalkaloids (Oda *et al*, 2002) and bacterial virulence factors (Deng *et al*, 2000). In plants L-rhamnose-containing terpenyl glycosides are important aroma precursors (Günata *et al.*, 1985) and may also play a protective role against the toxicity of free aglycons; L-rhamnose-containing flavonoid glycosides have antioxidant and anti-inflammatory activities (Benavente-García *et al*, 1997).

α-L-rhamnosidases (EC 3.2.1.40) and β-L-rhamnosidases (EC 3.2.1.43) catalyse the hydrolysis of terminal, non-reducing L-rhamnose residues in α- and β-L-rhamnosides respectively. In contrast, endorhamnosidases (EC 3.2.1.-) act by cleaving specific linkages between internal rhamnose residues in rhamnosides. α-L-rhamnosidases have been found in many micro-organisms and in some plant and animal tissues (see below), whereas β-L-rhamnosidase has only been described

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 117-140

in *Klebsiella aerogenes* (Barker *et al*., 1965). Endorhamnosidases seem to be restricted to bacteriophages (Steinbacher *et al*., 1994; Chua *et al*., 1999). This chapter will mainly focus on the α-L-rhamnosidases (αRHAs) as a group of hydrolytic enzymes having crucial biological functions and important potential biotechnological applications.

In 1991 a classification of glycoside hydrolases based on amino acid sequence similarities was introduced (Henrissat, 1991). This classification, which is regularly updated, now comprises more than 100 sequence-based families (URL://http://afmb.cnrs-mrs.fr/CAZY/). The αRHAs, with the exception of that of *Sphingomonas paucimobilis* that has recently been assigned to family 106, belong to family 78 (Coutinho and Henrissat, 1999). Currently, 61 different putative αRHAs from bacteria, yeasts and moulds are known. Genome-sequencing projects, particularly those focussing on bacterial, fungal and plant genomes, are beginning to generate large numbers of potential αRHA sequences, and a very recent search of the NCBI database (http://www.ncbi.nlm.nih.gov) yielded more than 100 entries for putative αRHA encoding genes. Endorhamnosidases, with the exception of that of bacteriophage Sf6 which has not been classified, are placed in family 90.

αRHAs are inverting glycoside hydrolases. Enzymatic hydrolysis of the glycosidic bond takes place via general acid catalysis that requires two critical residues: a proton donor and a nucleophile/base (Davies and Henrissat, 1995). These residues are as yet unidentified in αRHAs. In general, inverting glycosyl hydrolases (GH) typically employ two side chain carboxylates (supplied by Asp or Glu) in the active site to effect catalysis.

αRHAs have recently been the focus of several research initiatives because of their key roles in fundamental biological processes (*e.g.* detoxification mechanisms, symbiosis) and utility in biotechnological applications (*e.g.* elucidation of the structures of biologically important glycosides, biomass conversion, beverage quality enhancement and the manufacture of hydrolysis products from natural glycosides).

2. BIOCHEMICAL PROPERTIES AND STRUCTURE

2.1. Sources and Biochemical Characteristics of α-L-Rhamnosidases

Although micro-organisms are the main sources of αRHAs, these enzymes have also been found in animal tissues such as the liver of the marine gastropod *Turbo cornutus* (Kurosawa *et al*., 1973) and pig (Qian *et al*., 2005), and plants such as *Rhamnus daurica* (Suzuki, 1962) and *Fagopyrum esculentum* (Bourbouze *et al.*, 1976). Given the sparseness of reports on animal and plant αRHAs, data presented hereafter are mainly focused on microbial enzymes.

αRHAs seem to be common in filamentous fungi as revealed by recent screenings that have identified different strains of *Acremonium*, *Aspergillus*, *Circinella*, *Eurotium*, *Fusarium*, *Mortierella*, *Mucor*, *Penicillium*, *Rhizopus*, *Talaromyces* and *Trichoderma* as αRHA producers (Scaroni *et al*., 2002; Monti *et al*., 2004). The presence of αRHA activity in phytopathogenic fungi was first described

in *Corticium rolfsii* (Kaji and Ichimi, 1973) and more recently in *Stagonospora avenae*, a cereal pathogen (Morrissey *et al*., 2000; Hughes *et al*., 2004). In addition, highly specific αRHAs from *Absidia* sp. (Yu *et al*., 2002, 2004) and *Plectosphaerella cucumerina* (Oda *et al*., 2002) which are able to degrade different saponins have been purified and characterized. Nevertheless, biochemical characterization of fungal αRHAs has been carried out mainly on enzymes purified from culture filtrates and on commercial enzyme preparations from *Aspergillus* species such as *A. terreus* (Gallego *et al*., 1996; 2001), *A. nidulans* (Orejas *et al*., 1999; Manzanares *et al*., 2000), *A. niger* (Kurosawa *et al*., 1973; Manzanares *et al*., 1997) and *A. aculetaus* (Mutter *et al*., 1994; Manzanares *et al*., 2001, 2003), and also on commercial preparations of *Penicillium* species (Romero *et al*., 1985; Young *et al*., 1989).

Regarding yeasts, low levels of αRHA activity have been found in screenings performed on oenological yeast strains (Miklosy and Polos, 1995; McMahon *et al*., 1999; Rodríguez *et al*., 2004), and so far one intracellular αRHA has been purified and characterized from *Pichia angusta* (Yanai and Sato, 2000).

αRHA-producing bacteria have been described in human faecal flora belonging to the genus *Bacteroides* (Bokkenheuser *et al*., 1987; Jang and Kim, 1996), in a *Bacillus* sp. and in *S. paucimobilis* (originally designated *Pseudomonas paucimobilis*) strains isolated from soil (Hashimoto *et al*., 1999, 2003; Miake *et al*., 2000; Miyata *et al*., 2005), as well as in the thermophilic bacteria *Clostridium stercorarium* (Zverlov *et al*., 2000) and *Thermomicrobia* sp. (Birgisson *et al*., 2004). αRHA activity has also recently been described in wine strains of *Oenococcus oeni* (Grimaldi *et al*., 2005).

αRHAs from various microbial sources have been purified and some of the general properties of these enzymes are summarized in Table 1. Data presented for fungal proteins correspond to αRHAs purified from wild strains whereas the characterizations of *Bacillus* sp., *C. stercorarium* and *Thermomicrobia* sp. enzymes were carried out on recombinant counterparts. The main difference found between fungal and bacterial enzymes concerns their optimal pH. With the exception of *P. angusta* αRHA, the fungal enzymes show acidic pH optima compared to bacterial αRHAs for which neutral and alkaline pH optima have been found. This characteristic suggests different potential applications for fungal and bacterial enzymes, making fungal αRHAs more suitable for use in processes operating at low pH such as winemaking (Manzanares *et al*., 2003) and citrus juice processing (Puri *et al*., 1996). The bacterial enzymes would be useful in processes requiring good activity in basic solutions, such as the production of L-rhamnose by hesperidin hydrolysis given that the solubility of this flavonoid glycoside increases at high pH (Scaroni *et al*., 2002). Differences in αRHA cellular location have been also found, suggesting different *in vivo* roles for fungal and bacterial αRHAs.

Table 2 shows the catalytic properties of some purified microbial αRHAs. Studies of substrate specificity and the observed inhibition of activity by L-rhamnose clearly demonstrate the specificity of αRHAs for rhamnopyranoside residues. By contrast, the type of linkage present at C1 of α-L-rhamnose has only a minor effect on

Table 1. General properties of some purified microbial α-L-rhamnosidases

Micro-organism	Original strain or overexpression host	Inducer	pH optimum	Temp. optimum (°C)	Molecular weight (kDa)	Native form	pI	Cellular location	Reference
Fungi									
Absidia sp.	Original strain								
Absidia sp. 39		Ginseng	5	40	53	n.d.	n.d.	Extracellular	Yu
Absidia sp. 90		*Gynostemma pentaphyllum*	5	50	68	n.d.	n.d.	Extracellular	*et al.,* 2002, Yu *et al.,* 2004
A. aculeatus	Original strain	Hesperidin	4.5–5	n.d.	92 (RhaA)[1]	n.d.	6.2	Extracellular	Manzanares
			4.5–5	n.d.	85 (RhaB)[1]	n.d.	5.2–5.9	Extracellular	*et al.,* 2001, 2003
A. nidulans	Original strain	Rhamnose	4.5–6	60	102[1]	n.d.	5	Extracellular	Manzanares *et al.,* 2000
A. niger	Original strain	Hesperidin	4.5	65	85[1]	n.d.	4.5–5.2	Commercial prep	Kurosawa *et al.,* 1973; Manzanares *et al.,* 1997
A. terreus	Original strain	Rhamnose	4	44	96	n.d.	4.6	Extracellular	Gallego *et al.,* 2001
Penicillium sp.	Original strain	Naringin	3.5	57	90[1]	n.d.	n.d.	Commercial prep	Romero *et al.,* 1985; Young *et al.,* 1989

P. angusta	Original strain	Rhamnose	6	40	90	Monomer	4.9	Intracellular	Yanai and Sato, 2000
S. avenae	Original strain	Starch	n.d.	n.d.	110	n.d.	n.d.	Extracellular	Hughes *et al.,* 2004
Bacteria									
Bacillus sp.[2]	Original strain	Gellan	6.5–7	40	98 (RhaA)[3]	Pentamer	n.d.	Intracellular	Hashimoto
	E. coli		6.5–7	40	106 (RhaB)	Monomer	n.d.	Intracellular	*et al.,* 2003
Bacteroides sp.	Original strain		7		120[3]	Dimer	4.2	Intracellular	Jang and Kim, 1996
C. stercorarium[2]	*E. coli*		7.5	60	95 (RamA)[3]	Dimer	n.d.	Intracellular	Zverlov *et al.,* 2000
S. paucimobilis	Original strain	Rhamnose	7.8	45	112 (RhaM)	Monomer	7.1	Intracellular	Miake *et al.,* 2000
Thermomicrobia[2]	*E. coli*		7.9	70	104 (RhmA)[3]	Dimer	4.6	Intracellular	Birgisson *et al.,* 2005
			5–6.9	70	107 (RhmB)[3]	Dimer	4.5	Intracellular	

n. d.: not determined. [1]Molecular weight of glycosylated protein. [2]Data from recombinant enzymes. [3]Monomeric form molecular weight.

Table 2. Catalytic properties of some microbial purified α-L-rhamnosidases

Micro-organism	K_m (mM)[1]	Substrate specificity	Linkage hydrolysed	Inhibitors	Reference
Fungi					
Absidia sp.					
Absidia sp. 39	n.d.	Ginsenoside Rg2	α − 1, 2	n.d.	Yu *et al.,* 2002, 2004
Absidia sp. 90	n.d.	Gipenoside-5	α − 1, 6	n.d.	
A. aculeatus					Manzanares
RhaA	2.8	Hesperidin, naringin, rutin, monoterpenyl-glycosides	α − 1, 2; α − 1, 6	Rhamnose	*et al.,* 2001, 2003
RhaB	0.30	Hesperidin, naringin, rutin	α − 1, 2; α − 1, 6	Rhamnose	
A. nidulans	0.27	Hesperidin, naringin, rutin	α − 1, 2; α − 1, 6	Rhamnose, Mg^{2+}	Manzanares *et al.,* 2000
A. niger	2.9	Hesperidin, naringin, rutin geranyl-β-D-rutinoside 2-phenylethyl-β-D-rutinoside	α − 1, 2; α − 1, 6	Rhamnose	Manzanares *et al.,* 1997
A. terreus	0.17	Naringin, rutin	α − 1, 2; α − 1, 6	Rhamnose, ethanol, glucose, Hg^{2+}, Cd^{2+}	Gallego *et al.,* 2001
Penicillium sp.	1.52	Naringin	α − 1, 2	Rhamnose, glucose	Romero *et al.,* 1985
P. angusta	n.d.	Hesperidin, naringin, quercitrin, rutin, aroma precursors from grape juice	α − 1, 2; α − 1, 6; α1 to aglycon	Rhamnose, Cu^{2+}, Hg^{2+}, *p*-chloromercuribenzoate	Yanai and Sato, 2000

S. avenae	0.091	Avenacoside A and B	$\alpha-1,4$	n.d.	Hughes *et al.,* 2004
Bacteria					
Bacillus sp.	0.119	Naringin, gellan disaccharide	$\alpha-1,2$; $\alpha-1,3$	Rhamnose, Hg^{2+}	Hashimoto
RhaA	0.282	Naringin, gellan disaccharide	$\alpha-1,2$; $\alpha-1,3$	Rhamnose, Hg^{2+}, Cu^{2+}, Fe^{3+},	*et al.,* 1999, 2003
RhaB				glucose, 6-deoxyglucose	
Bacteroides sp.	0.29	Hesperidin, naringin, neohesperidin, poncirin, quercitrin, robinin, rutin, saikosaponin C		Rhamnose, fucose, saccharic acid, 1,4-lactone, Pb^{2+}, *p*-chloromercuriphenylsulfonic acid	Bokkenheuser *et al.,* 1987; Jang and Kim, 1996
C. stercorarium	n.d.	Hesperidin, naringin	$\alpha-1,2$; $\alpha-1,6$	Hg^{2+}, Cu^{2+}, Zn^{2+}, SDS, *p*-chloromercuribenzoate	Zverlov *et al.,* 2000
S. paucimobilis	1.18	Hesperidin, naringin, proscillaridin A, quercitrin, rutin, saikosaponin C, rhamnosyl-disaccharides	$\alpha-1,2$; $\alpha-1,3$; $\alpha-1,4$; $\alpha-1,6$; $\alpha1$ to aglycon	Cu^{2+}, Pb^{2+}, Cd^{2+}, Zn^{2+}, Ba^{2+}	Miake *et al.,* 2000
Thermomicrobia					Birgisson *et al.,* 2004
RhmA	0.46	Naringin, hesperidin, rutin	$\alpha-1,2$; $\alpha-1,6$	Rhamnose, glucose, ethanol	
RhmB	0.66	Naringin, hesperidin	$\alpha-1,2$; $\alpha-1,6$	Rhamnose, glucose, ethanol, Zn^{2+}	

[1]using pNPR as substrate with the exception of *S. avenae* α-L-rhamnosidase K_m determined on 26-desglucoavenacoside.
n.d.: not determined

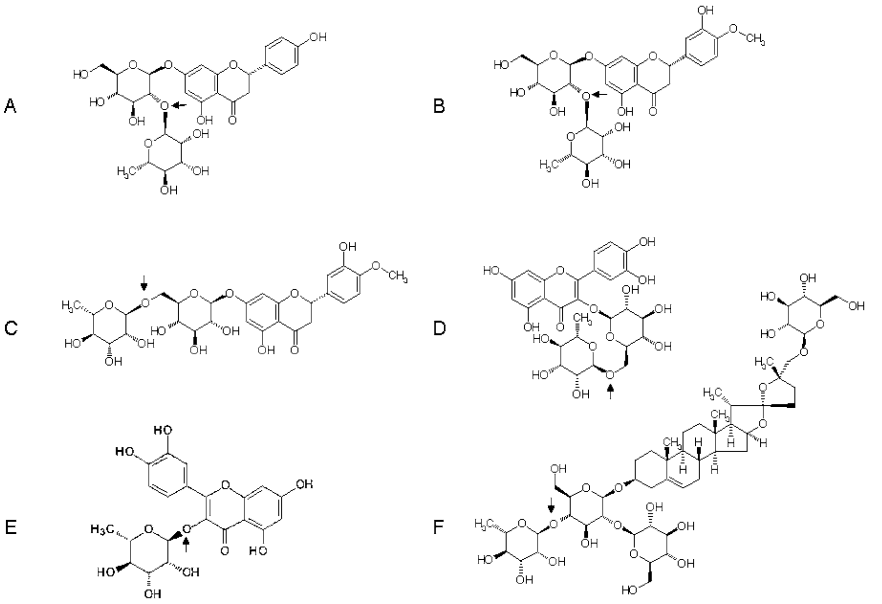


Figure 1. Structure of some substrates hydrolysed by αRHAs. A: naringin; B: neohesperidin; C: hesperidin; D: rutin; E: quercitrin; F: avenacoside A. The arrows indicate the possible linkages hydrolysed by αRHAs

activity since αRHAs are able to hydrolyse various substrates. These enzymes are able to cleave both glycosidic bonds and aglycon-saccharide bonds (Fig. 1). As regards glycosidic bonds, αRHAs hydrolyse substrates in which the L-rhamnose residue is linked to a β-glucosidic residue through either i) an α-1,2 linkage such as that present in the flavonoids naringin, neohesperidin and poncirin and in ginseng saponins, ii) an α-1,3 linkage such as that in the gellan disaccharide α-L-rham-(1-3)-β-D-glc, iii) an α-1,4-linkage such as that in oat saponins, or iv) an α-1,6 linkage, as found in the flavonoids hesperidin and rutin, in the grape glycosides geranyl- and 2-phenylethyl-rutinosides, and in the saponin gypenoside-5. αRHAs are also able to hydrolyse substrates in which L-rhamnose is directly linked to an aglycon, as is the case in the flavonoids quercitrin and robinin, and in the aryl-rhamnoside ρ-nitrophenyl-α-L-rhamnopyranoside (pNPR) the model substrate for assaying rhamnohydrolase activity. The latter has been commonly used for determining αRHA K_m values, which can range from 0.119 to 2.9 mM (see Table 2). K_m values for hesperidin (0.06 mM), rutin (0.13 mM), naringin (0.17 mM) and quercitrin (0.18 mM) for *S. paucimobilis* αRHA have also been determined (Miake *et al.*, 2000). A K_m value of 7.0 mM for naringin has been described for *Penicillium* sp. αRHA (Romero *et al*, 1985).

Differences in aglycon structure may explain the differences observed in the hydrolysis of glycosides having the same linkages. For instance the reason why

rutin and hesperidin (both of which have an α-1,6 linkage) are so differently hydrolysed by *Aspergillus* αRHAs may be explained by steric hindrance due to the attachment of the diglycoside to the aglycon molecule via C7 in hesperidin whereas the attachment is via C3 in rutin. Differences in the hydrolysis of substrates in which L-rhamnose is directly linked to the aglycon have also been described, and in general αRHAs show a clear preference for pNPR in comparison to quercitrin or robinin. However, αRHAs able to hydrolyse different saponins such as those from *Absidia* species (Yu *et al*., 2002, 2004) and *S. avenae* (Hughes *et al*., 2004) hardly hydrolyse pNPR.

2.2. Structure

Compared with the number of αRHAs that have been purified and characterized, the number of genes known that encode αRHAs is much lower. Cloning has been performed mainly by two methods: i) construction of a library followed by selection of clones by screening for αRHA activity using pNPR or 4-methylumbelliferyl-α-L-rhamnoside plate assays, and/or ii) the construction of a library followed by selection of clones by screening for αRHA production with polyclonal antibodies. The first αRHA encoding gene (*ramA*), isolated from the thermophilic anaerobic bacterium *C. stercorarium* (Zverlov *et al*., 2000), consists of 873 or 874 codons and codes for a protein with a predicted molecular mass of 100 kDa and a global secondary structure, as detected by circular dichroism spectroscopy, consisting of 27% α-helices and 50% β-sheets. Soon thereafter the genes *rhaA* and *rhaB* encoding two *A. aculeatus* αRHAs were cloned (Manzanares *et al*., 2001). Conceptual translation of the cDNAs corresponding to *rhaA* and *rhaB* yields primary structures of 660 and 597 amino acids (71 and 64 kDa, respectively). However, the N-terminal amino acid sequences actually determined for RhaA and RhaB correspond to positions 20 and 17, hence the derived molecular masses of the mature proteins are 69 and 62 kDa.

To date, only eight αRHA genes have been cloned and heterologously expressed: those from *C. stercorarium* and *A. aculeatus* (Zverlov *et al*., 2000; Manzanares *et al*., 2001), two genes (*rhaA* and *rhaB*) from *Bacillus* sp. (Hashimoto *et al*., 2003), two genes (*rhmA* and *rhmB*) from *Thermomicrobia* sp. (Birgisson *et al*., 2004) and one gene (*rhaM*) from *S. paucimobilis* (Miyata *et al*., 2005).

The value of genome sequences for the rapid identification of candidate αRHA genes is obvious. Hypothetical proteins of unknown function found by such analyses are putative novel αRHAs. For example, two sequences from *Lactobacillus acidophilus* and *Aspergillus fumigatus* exhibited sequence similarities with the enzymes RamA of *C. stercorarium* and RhaA of *A. aculeatus*, respectively.

Whereas electrophoretic and chromatographic detection of αRHAs indicate the presence of only one or two isoenzymes in a given organism, data from genome sequencing projects suggest the presence of many different isoforms. We have previously reported (Orejas *et al*., 1999; Manzanares *et al*., 2000) the production of an apparently unique extracellular αRHA in *A. nidulans* with a molecular mass of 102 kDa. However, *in silico* analysis of the *A. nidulans* genome reveals the

presence of 8 ORFs encoding hypothetical αRHAs ranging from 63 to 253 kDa (family GH 78 in the CAZY database; de Vries *et al*, 2005), the functions of which still have to be demonstrated. The presence of various 'isoforms' of αRHA within the filamentous fungi was also suggested by Monti *et al*. (2004) based on inducer and substrate specificity data.

To date no three-dimensional structure of αRHA has been published.

2.3. Sequence Comparisons

Some micro-organisms have already been reported that produce two αRHAs. Whereas in the case of *A. aculeatus* the primary structures of RhaA and RhaB are highly homologous (60% identity) (Manzanares *et al*, 2001), those produced by *Bacillus* sp. strain GL1 and those produced by *Thermomicrobia* sp. strain PRI-1686 have quite different primary structures (only 23% identity) (Hashimoto *et al*, 2003; Birgisson *et al*, 2004).

Using the FASTA program (Pearson, 1990), RhaA of *Bacillus* sp. was found to exhibit significant homology with RamA of *C. stercorarium* (41% identity in 848 aa overlap) but no homology with either RhaA or RhaB of *A. aculeatus* (Hashimoto *et al*, 2003). Despite the observed similarity between *Bacillus* RhaA and *C. stercorarium* RamA, these enzymes are distinct in their quaternary structures (pentameric *versus* homodimeric, see Table 1). By contrast, *Bacillus* RhaB exhibits ~24% identity with RhaA and RhaB of *A. aculeatus* (Hashimoto *et al*, 2003) but differs significantly from the fungal enzymes since the *A. aculeatus* αRHAs have considerably lower molecular masses as calculated by conceptual translation of the ORFs (see above). Using the BLAST program (Altschul *et al*, 1997) it has been shown that *Thermomicrobia* RhmA and RhmB display greatest similarity to the isozymes RhaA and RhaB of *Bacillus* sp.: RhmA has 41% identity with RhaA and RhmB has 50% identity with RhaB.

The *Sphingomonas* enzyme RhaM has no similarity to other known αRHAs, and Miyata *et al*. (2005) have suggested that this protein could be a member of a new bacterial subfamily within glycoside hydrolase family 78. In BLAST analysis, RhaM showed 58% identity and 72% homology (in 1113 aa overlaps) with the hypothetical protein Saro 02001624 of *Novosphingobium aromaticivorans*. This hypothetical protein, predicted from complete genome sequence analysis, may thus also have αRHA activity.

In order to locate putative catalytic sites in αRHAs, an amino acid sequence alignment using the ClustalW program (Thompson *et al*, 1994; Higgings *et al*, 1996) was performed with those members of the GH78 family for which the corresponding encoding genes have been functionally characterised (Fig. 2). Since glycosyl hydrolases require the presence of two or more carboxylic acid moieties for their function, the aligned sequences were examined for conserved Asp and Glu residues. Two fully conserved and four well conserved carboxylic amino acid residues were found in the alignment, all of which are possible candidates for playing an important role in αRHA catalytic function. Site-directed mutagenesis

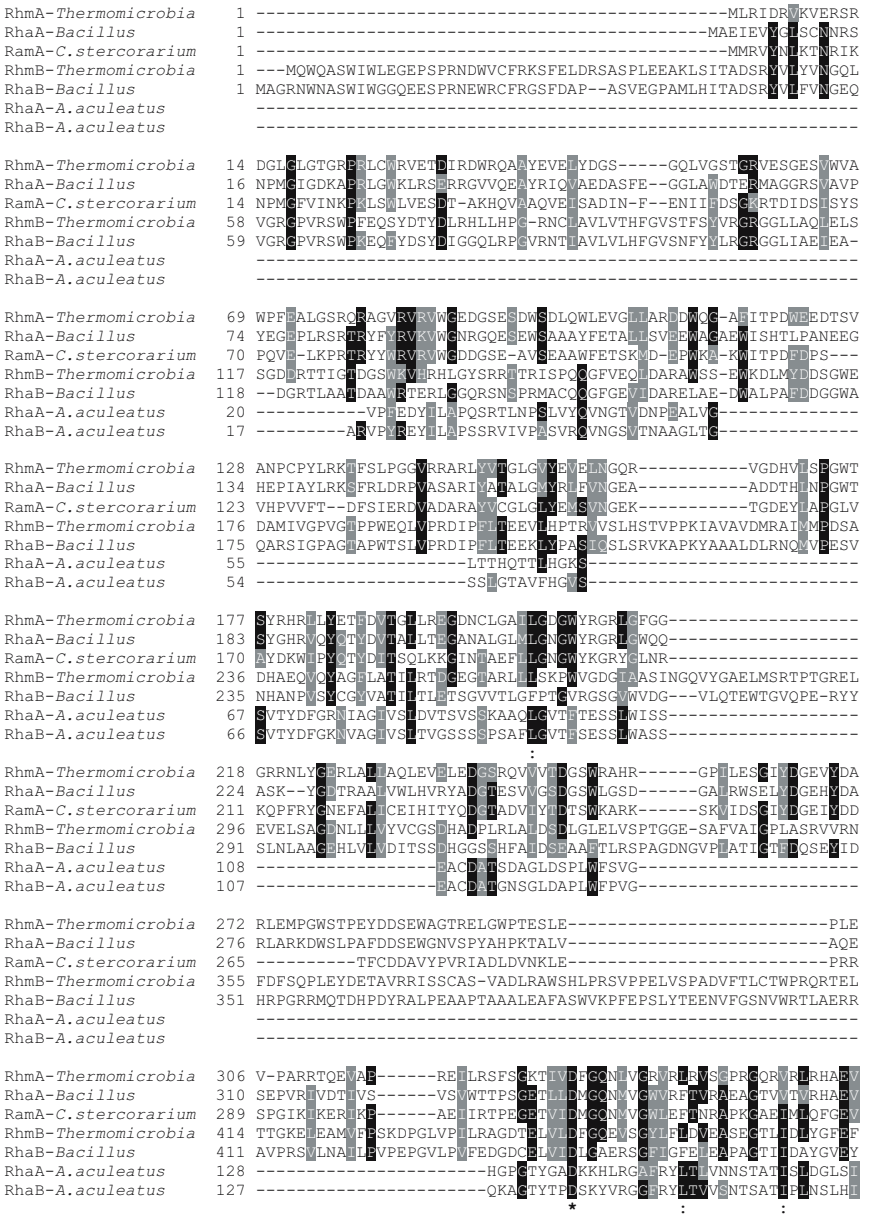


Figure 2. (Continued)

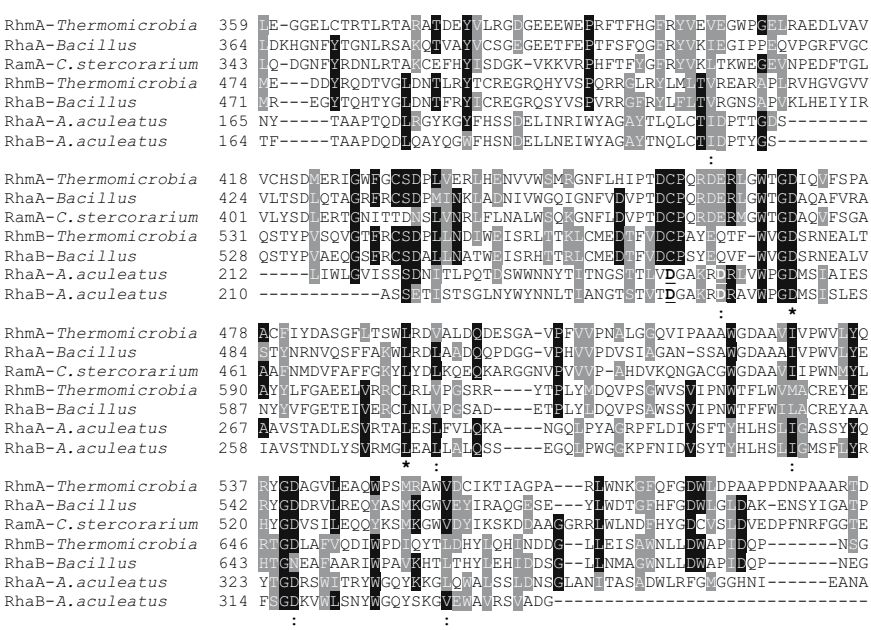


Figure 2. (Continued)

Figure 2. (Continued) Multiple sequence alignment of some members of glycosyl hydrolase family GH78. The alignment was performed using ClustalW (v 1.82) (Thompson *et al*., 1994) and shading was performed with Box-shade. The sequences aligned are listed below followed by their GenBank accession numbers in parentheses. RamA of *C. stercorarium* (AJ238748), RhaA (AF284761) and RhaB (AF284762) from *A. aculeatus*, RhaA (AB046705) and RhaB (AB046706) from *Bacillus* sp. GL1, and RhmA (AY505013) and RhmB (AY505014) from *Thermomicrobia* sp. PRI-1686. In the case of the *A. aculeatus* proteins, the putative signal peptides were not included. Identical residues are depicted on a black background whereas similar residues are shown on a grey background. Conserved carboxylic amino acid residues and hydrophobic amino acids are marked with * and : respectively

of putative key residues involved in catalysis will indicate their *in vivo* relevance and potentially provide a means to modify enzyme characteristics. A phylogenetic tree (Fig. 3) based on primary sequence homology was also constructed for the enzymes analysed in Fig. 2 along with RhaM of *S. paucimobilis*. The pattern obtained suggests that these enzymes evolved from a common ancestor into at least 3 distinct clusters.

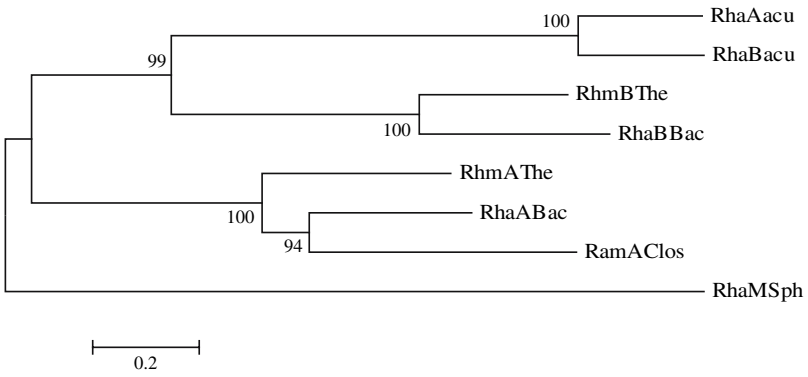


Figure 3. Phylogenetic analysis of sequence similarities between αRHAs. MEGA 3 software (Kumar *et al.*, 2004) was used to carry out the analysis of those αRHAs included in Fig. 2 as well as the mature form of *S. paucimobilis* RhaM (GenBank accession no. AB080801). Bootstrap values are adjacent to each internal node, representing the percentages of 1000 bootstrap replicates. The scale represents amino acid replacements per residue

3. BIOLOGICAL FUNCTION

Neither the reason for the redundancy of the genes encoding αRHAs nor the *in vivo* function of αRHAs are completely understood. Speculation on the biological role of these enzymes is mainly based on studies of the catalytic properties and substrate specificities of purified αRHAs (Table 2). Recently, the design of potent αRHA inhibitors has been suggested as a tool to investigate the biological function of the enzyme (Kim *et al*., 2005).

The low specific activities found on the natural substrates tested and the fact that most of the αRHAs characterized show a clear preference for pNPR has led some authors to suggest that the preferred natural substrate for these glycosidases is still unknown (Zverlov *et al*., 2000; Manzanares *et al*., 2001). Nevertheless, and as described below, several *in vivo* functions including heteropolysaccharide and flavonoid metabolism as well as detoxification processes have been described.

It seems clear that at least RhaB from *Bacillus* sp. GL1 is indispensably required for the complete metabolism of gellan and gellan-related polysaccharides (sphingans), heteropolysaccharides produced by different species belonging to the genera *Sphingomonas* (Hashimoto *et al*., 1999, 2003). Given the co-isolation of *Bacillus* sp. GL1 and *Sphingomonas* sp. R1 from soil samples, it has been suggested that these micro-organisms could be symbiotic and that due to its gellan degrading ability *Bacillus* sp. GL1 would be able to survive in a capsule-like sphingan biofilm formed by *Sphingomonas* sp. R1. The capsule would thus function as a barrier to inhibit diffusion of the low molecular weight depolymerization products necessary for their growth.

It has also been suggested that αRHA may be involved in flagellum formation based on results showing that *S. paucimobilis* αRHA production is induced by L-rhamnose and that only in the presence of this sugar does *S. paucimobilis* form flagella (Miake *et al*., 1995).

Glycosidases produced by the human intestinal microflora are known to participate in the degradation of dietary flavonoids (Bokkenheuser *et al*., 1987; Jang and Kim, 1996), the first step being hydrolysis to yield the aglycon. This conversion probably takes place in the lower part of the ileum and the caecum. The aglycon molecules are then either further bacterially metabolised or absorbed into the enterohepatic system. Studies on the metabolism of some ginseng glycosides have suggested a key role for αRHAs and β-glucosidases in the manifestation of the pharmacological properties of ginseng such as its oestrogenic effect (Bae *et al*., 2005).

The αRHAs of phytopathogenic fungi seem to be involved in overcoming saponin-mediated plant defences. Plant saponins are glycosylated compounds that can repress the growth of fungi (Osbourn, 1996a). Successful plant pathogens avoid the antifungal properties of saponins by modifying their membrane composition and/or by detoxifying saponins via hydrolytic removal of their sugar moieties (Osbourn, 1996b). The αRHA of *S. avenae*, a fungus able to infect oat leaves (Morrissey *et al*., 2000), is involved in the latter strategy. In response to pathogen attack, biologically inactive plant avenacosides saponins are converted

into 26-desglucoavenacosides which possess antifungal activity. These molecules are comprised of a steroidal backbone linked to a branched sugar chain consisting of one α-L-rhamnose and two or three β-D-glucose residues. Isolates of the fungus that are pathogenic to oats are capable of sequentially hydrolysing these sugar residues. Degradation is initiated by removal of the L-rhamnose which abolishes the antifungal activity. A similar function has been described for the αRHAs of several filamentous fungi (*Cladosporium cladosporioides*, *Penicillium* sp. and *Plectosphaerella cucumerina*) which are able to grow on potato sprouts despite the accumulation by the latter of the antifungal α-chaconine (Oda *et al*., 2002). The hydrolysis of one of the two L-rhamnose residues in α-chaconine seems to be the first step in the latter's detoxification.

4. POTENTIAL INDUSTRIAL APPLICATIONS OF α-L-RHAMNOSIDASES

Nowadays enzyme technology presents an alternative to chemical processes, reducing both energy and material consumption and minimizing the generation of wastes and emissions. In this context, αRHAs have been demonstrated to be of biotechnological utility with possible industrial applications. These applications, directed to the food, pharmaceutical and chemical industries, are mainly based on αRHA hydrolytic activity although some applications based on synthetic activity have also been described (Table 3).

4.1. Food Industry

The main use of αRHA in the food industry is for beverage quality enhancement (debittering, liberation of aromas and bioactive compounds) and the production of food additives.

Biotechnological approaches for the debittering of citrus juices are focused on the development of enzymes capable of hydrolysing naringin and limonin, the two major contributors to bitterness in processed citrus juices. αRHA is involved together with β-D-glucosidase in the stepwise hydrolysis of naringin (see Fig. 1). Both activities, collectively termed 'naringinase', work sequentially. αRHA splits naringin into rhamnose and prunin, and ß-D-glucosidase splits prunin into glucose and naringenin. Naringenin bitterness is only one third that of naringin, and prunin is less bitter than naringenin. In fact, only the first hydrolysing activity, αRHA, is essential. The feasibility of the enzymatic approach to debittering has been shown with both soluble and immobilized naringinase (Norouzian *et al*., 1999; Prakash *et al*., 2002; Puri *et al*., 2005). Active packaging has also been described as an alternative method to reduce the naringin content of citrus juices during storage and transport by means of direct interaction with the product. The system uses an 'active' film comprising a cross-linked matrix in which naringinase is completely immobilized. Various data suggest that the active package developed

Table 3. Possible industrial applications of α-L-rhamnosidases in the food, pharmaceutical and chemical industries

	Application	Reference
Food industry		
Beverages	Debittering of citrus juice by hydrolysis of naringin either directly or by active packaging	Soares and Hotchkiss, 1998; Norouzian *et al.,* 1999; Prakash *et al.,* 2002; Del Nobile *et al.,* 2003; Puri *et al.,* 2005
	Improvement of wine aroma by hydrolysis of aromatic precursors	Günata, 2003; Manzanares *et al.,* 2003
	Enhancement of fruit juice functional properties by increasing flavonoid bioavailability	Hollman *et al.,* 1999; González-Barrio *et al.,* 2004
Additives	Production of natural sweetener precursor by hydrolysis of hesperidin	Sánchez *et al.,* 1987
	Obtention of new ingredients from biopolymers to enhance food rheological properties	Jansson and Lindberg, 1983; Hashimoto *et al.,* 1999; Giavasis *et al.,* 2000
Pharmaceutical industry	Production of compounds with enhanced pharmacological properties by enzymatic hydrolysis of different rhamnosides	Mimaki *et al.,* 1998; Acquati and Ponzone, 2000; Di Lazzaro *et al.,* 2001; Boyle *et al.,* 2003; Pisvejcová *et al.,* 2003; Yu *et al.,* 2004
	Obtention of novel glycopeptide antibiotics	Takatsu *et al.,* 1987a,b
Chemical industry	Design of low cost biotechnological processes for the production of pure compounds	Matsumoto *et al.,* 2002; Trummler *et al.,* 2003; Chang and Muir, 2004
	Production of L-rhamnose as a precursor for industrial use or as a chiral compound for chemical synthesis	Martearena *et al.,* 2003

can be successfully used to improve the sensory properties of grapefruit juices. (Soares and Hotchkiss, 1998; Del Nobile *et al*., 2003).

Since the demonstration that the aromatic components of certain grape varieties are present in the grape berry both in free form and also bound to sugars in the form of glycosides, the usefulness of glycosidases for the release of varietal aromas from precursor compounds during winemaking has been investigated. The bound aroma fraction comprises glucosides and disaccharide glycosides such as 6-O-α-L-arabinofuranosyl-β-D-glucopyranosides, 6-O-α-L-rhamnopyranosyl-β-D-glucopyranosides and 6-O-β-D-apiofuranosyl-β-D-glucopyranosides. Compounds such as terpenols, terpene diols, 2-phenylethanol, benzyl alcohol and C_{13} norisoprenoids have been shown to be aglycons of such glycosides. Enzymatic hydrolysis of the aroma precursor compounds requires two sequential reactions: first, an αRHA, an α-L-arabinofuranosidase, or a ß-D-apiosidase cleaves the α-1,6 glycosidic linkage; subsequently the aroma/flavour compounds are liberated from the monoglucosides by the action of a ß-D-glucosidase. Since grape and yeast glycosidases seem to be insufficient to process aromatic precursors completely during winemaking, the addition of exogenous glycosidases during or after the fermentation is now common practise in wineries. αRHA is a component of these commercial enzymatic preparations and its key role in aroma release has been established. Contrary to that which is described for debittering processes, in order to develop the aromatic potential of a wine to the full all glycosidase activities are essential (for a review see Günata, 2003). Besides exogenous enzyme addition, it is possible to achieve increases in the content of volatile compounds during vinification by using recombinant wine yeast strains expressing such hydrolytic activities. A genetically modified industrial wine yeast strain expressing the *A. aculeatus rha*A gene has been constructed and wines produced in microvinifications conducted using a combination of this strain together with another strain expressing a ß-D-glucosidase showed increased content mainly of the aromatic compound linalool (Manzanares *et al.*, 2003).

Although the use of glycosidases to release flavour compounds from glycosidic precursors was initially examined in wines, fruit juice flavour may also be enhanced by αRHA application given the ubiquity of flavour glycoconjugates in fruits (Günata, 2003).

Since flavonoid glucosides have been reported to be more bioavailable than their rutinoside (glucose + rhamnose) counterparts (Hollman *et al*., 1999), both *A. aculeatus* αRHAs (RhaA and RhaB) have been used to produce functional beverages based on potentially increased flavonoid bioavailability (González-Barrio *et al.*, 2004). Incubation of blackcurrant juice, orange juice and green tea infusion with either RhaA or RhaB resulted in a decrease in the flavonoid rutinoside content (anthocyanins in blackcurrant juice, flavanones in orange juice and flavonols in green tea) and a concomitant increase in flavonoid glucosides.

With respect to the manufacture of food additives, αRHA could be used in the preparation of versatile food additives from biopolymers and in the production of sweeteners. Biopolymers contribute to food quality as gelling agents, thickeners,

stabilizers, lubricants, flocculants and flavour enhancers (Giavasis *et al*., 2000). Among the biopolymers, the bacterial exopolysaccharide gellan is principally composed of a tetrasaccharide repeating unit of one rhamnose, one glucuronic acid, and two glucose units. Various branching chains are attached to the repeating unit of the polysaccharide and determine the rheological properties of the polymer (Jansson and Lindberg, 1983). To date however the application of gellan gum has largely been limited due to its high viscosity. It has been proposed that enzymatic treatments, among which αRHA would play a key role, could be used to prepare low-viscosity and low molecular weight gellans having novel physiological and food-technological functions (Hashimoto *et al*., 1999). As regards sweeteners, another potential application of αRHA is based on its ability to cleave insoluble hesperidin (see Fig. 1) to rhamnose and hesperetin glucoside, the latter being a precursor in sweetener production (Sánchez *et al*., 1987).

4.2. Pharmaceutical Industry

The health-promoting activity of rhamnosides and their derivatives has opened a broad field for αRHA applications since their biological or pharmacological benefits have been observed to be inversely related to the amount of sugar residues present in the rhamnoside (Hollman *et al*., 1999; Chen *et al*., 2003). Studies investigating the impact of glycosidase treatments on the enhancement of the biological activities of rhamnosides are abundant in the literature, and several describe examples where the action of αRHA activity is crucial.

Extracts from the plant *Ruscus aculeatus* L. are known to possess various pharmacological properties including anti-inflammatory (Pisvejcová *et al*., 2003) and cytostatic (Mimaki *et al*., 1998) activities, and are also used in the treatment of chronic venous insufficiency (Boyle *et al*., 2003). The steroid saponins ruscin and ruscoside and their derivatives desglucoruscin, desglucodesrhamnoruscin and desglucoruscoside are the compounds that possess these properties. Biotechnological production of desglucoruscin and desglucodesrhamnoruscin (Fig. 4A) as well as new derivatives can be achieved employing an αRHA activity (Di Lazzaro *et al.*, 2001; Acquati and Ponzone, 2000).

Similar studies have been carried out with the triterpenoid saponin ginsenosides and gypenosides, the physiologically active compounds of some oriental herbs (Yu *et al.*, 2004). The removal of one rhamnose residue (Fig. 4B) converts gypenoside-5 into ginsenoside Rd that prevents kidney injury by anti-cancer drugs (Yokozawa and Owada, 1999).

Chloropolysporins A, B and C and their partially deglycosylated derivatives, new members of the glycopeptide antibiotic family, are active against Gram-positive bacteria, including clinically isolated methicillin-resistant *Staphylococci* and anaerobic enterobacteria. Derhamnosyl derivatives showing stronger activities than the parent compounds have been enzymatically obtained by treatment with αRHA (Takatsu *et al*., 1987a,b).

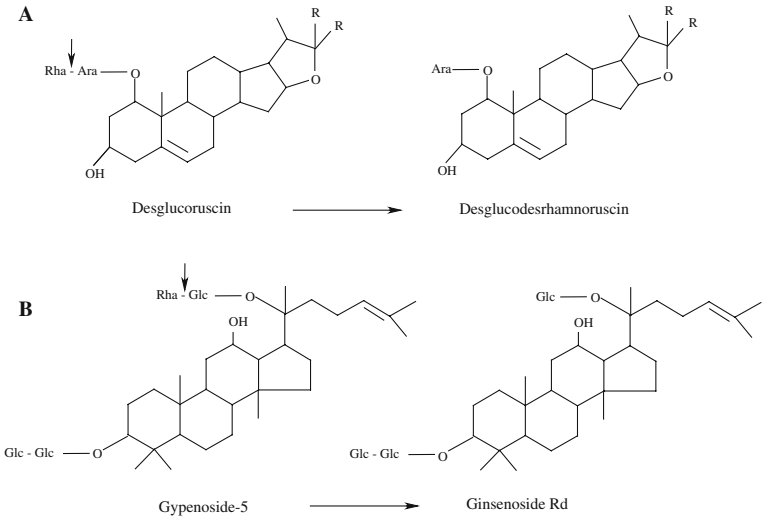


Figure 4. Chemical structures of desglucoruscin (A) and gypenoside-5 (B) and their αRHA conversion products. Arrows indicate the possible linkages hydrolysed by these enzymes

4.3. Chemical Industry

The application of αRHAs in the chemical industry is related to the design of low-cost processes for the production of valuable compounds. Among these, L-rhamnose has gained importance in recent years as both a precursor for the industrial production of aromatic compounds and flavours and as a chiral compound for chemical synthesis. As L-rhamnose is not biosynthesised as a free monomer it must be liberated from L-rhamnose-containing glycosides or polysaccharides, but this option is limited by the availability of suitable raw material. In this context, an integrated microbial/enzymatic process for the production of rhamnolipids and L-rhamnose from rapeseed oil has been developed (Trummler *et al.*, 2003). The process is aimed at improving the yield of L-rhamnose rather than rhamnolipids. This concept combines microbial rhamnolipid production by a *Pseudomonas* strain, with simultaneous enzymatic hydrolysis of rhamnolipid products in the same bioreactor.

Processes for the production of pure anthocyanidin glucosides from blackcurrant anthocyanidin rutinosides as well as the obtention of an isoquercitrin-enriched product from rutin by αRHA treatments have been recently patented (Matsumoto *et al.*, 2002; Chang and Muir, 2004).

Due to the importance of rhamnosides, glycosylation catalysed by αRHAs has been suggested as a way to produce pure rhamnosides in a single step. In this sense, a process has been proposed for the enzymatic synthesis of short chain length alkyl-α-L-rhamnosides using rhamnose or naringin as the glycosylation agents and water soluble alcohols as acceptors (Martearena *et al.*, 2003).

5. FUTURE PERSPECTIVES

In comparison to other glycosidases, knowledge of the molecular and structural characteristics of αRHAs as well as an understanding of their biological function is still scarce. Nevertheless the increasing importance of αRHA is reflected in the number of studies focused on possible applications. Since the first characterisation of an *α*RHA encoding gene in 2000 more αRHA genes are now available, and the application of DNA recombinant techniques for the overproduction of pure enzyme preparations and the modification of αRHA stability, selectivity or specificity are now feasible. These techniques will considerably extend the scope of potential applications and will convert αRHAs into an important industrial enzyme in the near future.

ACKNOWLEDGEMENTS

Work in the authors' laboratories was supported by CICYT grants AGL2002-01906, AGL2004-00978 and AGL2005-02542 (Spanish Ministry of Science and Education-FEDER).

REFERENCES

Acquati, W. and Ponzone, C. (2000) Preparation of desglucodesrhamnoruscin by hydrolysis of *Ruscus aculeatus* steroid glycosides by fermentation in the presence of *Aspergillus niger*. Patent number WO200073489-A.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acid Res. *25*, 3389–3402.

Bader, G., Wray, V., Just, U. and Hiller, K. (1998) Enzymatic hydrolysis of the cytotoxic glycoside virgaureasaponin 1. Phytochem. *49*, 153–156.

Bae, E.-A., Shin, J.-E. and Kim, D.-H. (2005) Metabolism of ginsenoside Re by human intestinal microflora and its estrogenic effect. Biol. Pharm. Bull. *28*, 1903–1908.

Bar-Peled, M., Lewinsohn, E., Fluhr, R. and Gressel, J. (1991) UDP-rhamnose-flavanone-7-O-glucoside-2"-O-rhamnosyltransferase-purification and characterization of an enzyme catalysing the production of bitter compounds in citrus. J. Biol. Chem. *226*, 20953–20959.

Barker, S.A., Somers, P.J. and Stacey, M. (1965) Arrangement of the L-rhamnose units in *Diplococcus pneumoniae* type II polysaccharide. Carbohydr. Res. *1*, 106–115.

Benavente-García, O., Castillo, J., Marín, F.R., Ortuño, A. and Del Río, J.A. (1997) Uses and properties of *Citrus* flavonoids. J. Agric. Food Chem. *45*, 4505–4515.

Birgisson, H., Hreggvidsson, G.O., Fridjónsson, O.H., Mort, A., Kristjánsson, J.K. and Mattiasson, B. (2004) Two new thermostable α-L-rhamnosidases from a novel thermophilic bacterium. Enzyme Microb. Technol. *34*, 561–571.

Bokkenheuser, V.D., Shackleton, C.H.L. and Winter, J. (1987) Hydrolysis of dietary flavonoid glycosides by strains of intestinal *Bacteroides* from humans. Biochem. J. *248*, 953–956.

Bourbouze, R., Percheron, F. and Courtois, J.E. (1976) α-L-Rhamnosidase from *Fagopyrum esculentum*: purification and some properties. Eur. J. Biochem. *63*, 331–337.

Boyle, P., Diehm, C. and Roberston, C. (2003) Meta-analysis of clinical trials of Cyclo 3 Fort in the treatment of chronic venous insufficiency. Int. Angiol. *22*, 250–262.

Chang, P.R. and Muir, A. (2004) Isoquercitrin-enriched composition prepared by incubating solution having rutin after adding naringinase, terminating the incubation, the proportion of isoquercitrin in composition being controlled by adjusting duration of incubation. Patent Number WO2004027074-A2.

Chen, Y.-C., Shen, S.-C. and Lin, H.-Y. (2003) Rutinoside at C7 attenuates the apoptosis-inducing activity of flavonoids. Biochem. Pharmacol. *66*, 1139–1150.

Chua, J.E.H., Manning, P.A. and Morona, R. (1999) The *Shigella flexneri* bacteriophage Sf6 tailspike protein (TSP)/endorhamnosidase is related to the bacteriophage P22 TSP and has a motif common to exo- and endoglycanases, and C-5 epimerases. Microbiol. *145*, 1649–1659.

Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-active Enzymes server at URL http://afmb.cnrs-mrs.fr/CAZY/.

Davies, G. and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. Structure *3*, 853–859.

Del Nobile, M.A., Piergiovanni, L., Buonocore, G.G., Fava, P., Puglisi, M.L. and Nicolais, L. (2003) Naringinase immobilization in polymeric films intended for food packaging applications. J. Food Sci. *68*, 2046–2049.

Deng, L., Kasper, D.L., Krick, T.P. and Wessels, M.R. (2000) Characterization of the linkage between the type III capsular polysaccharide and the bacterial cell wall of group B *Streptococcus*. J. Biol. Chem. *275*, 7497–7504.

Di Lazzaro, A., Morana, A., Schiraldi, C., Martino, A., Ponzone, C. and De Rosa, M. (2001) An enzymatic process for the production of the pharmacologically active glycoside desglucodesrhamnoruscin from *Ruscus aculeatus* L.J. Mol. Catal. B.-Enzym. *11*, 307–314.

Friedman, M. and McDonald, G.M. (1997) Potato glycoalkaloids: chemistry, analysis, safety, and plant physiology. Crit. Rev. Plant Sci. *16*, 55–132.

Gallego, M.V., Piñaga, F., Ramón, D. and Vallés, S. (1996) Production and characterization of an *Aspergillus terreus* α-L-rhamnosidase of oenological interest. Z. Lebensm.-Unters.-Forsch. *203*, 522–527.

Gallego, M.V., Piñaga, F., Ramón, D. and Vallés, S. (2001) Purification and characterization of an *Aspergillus terreus* α-L-rhamnosidase of interest in winemaking. J. Food Sci. *66*, 204–209.

Giavasis, I., Harvey, L.M. and McNeil, B. (2000) Gellan gum. Crit. Rev. Biotech. *20*, 177–211.

Grimaldi, A., Bartowsky, E. and Jiranek, V. (2005) A survey of glycosidase activities of commercial wine strains of *Oenococcus oeni*. Int. J. Food Microbiol. *105*, 233–244.

González-Barrio, R., Trindade, L.M., Manzanares, P., de Graaff, L. H., Tomás-Barberán, A. and Espín, J.C. (2004) Production of bioavailable flavonoid glucosides in fruit juices and green tea by use of fungal α-L-rhamnosidases. J. Agric. Food Chem. *52*, 6136–6142.

Günata, Z. (2003) Flavor enhancement in fruit juices and derived beverages by exogenous glycosidases and consequences of the use of enzyme preparations. In Handbook of Food Enzymology. Whitaker, J.R., Voragen, A.G.J. and Wong, D.W.S. eds. Marcel Dekker Inc., New York, pp 303–330.

Günata, Y.Z., Bayonove, C.L., Baumes, R.L. and Cordonnier, R.E. (1985) The aroma of grapes. I. Extraction and determination of free and glycosidically bound fractions of some grape aroma components. J. Chromatogr. *331*, 83–90.

Haruko, U. and Haruko, O. (1999) Glycobiology of the plant glycoprotein epitope: structure, immunogenicity and allergenicity of plant glycotopes. Trends Glycosci. Glycotech. *11*, 413–428.

Hashimoto, W., Miyake, O., Nankai, H. and Murata, K. (2003) Molecular identification of an α-L-rhamnosidase from *Bacillus* sp. strain GL1 as an enzyme involved in complete metabolism of gellan. Arch. Biochem. Biophys. *15*, 235–244.

Hashimoto, W. and Murata, K. (1998) α-L-Rhamnosidase of *Sphingomonas* sp. R1 producing an unusual exopolysaccharide of sphingan. Biosci. Biotechnol. Biochem. *62*, 1068–1074.

Hashimoto, W., Nankai, H., Sato, N., Kawai, S. and Murata, K. (1999) Characterization of α-L-rhamnosidase of *Bacillus* sp. GL1 responsible for the complete depolymerization of gellan. Arch. Biochem. Biophys. *368*, 56–60.

Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid similarities. Biochem. J. *280*, 309–316.

Higgings, D., Thompson, J. and Gibson, T. (1996) Using CLUSTAL for multiple sequence alignments. Methods Enzymol. *266*, 383–402.

Hollman, P.C., Bijsman, M.N., van Gameren, Y., Cnossen, E.P., de Vries, J.P. and Katan, M.B. (1999) The sugar moiety is a major determinant of the absorption of dietary flavonoid glycosides in man. Free Radical Res. *31*, 569–573.

Hughes, H.B., Morrissey, J.P. and Osbourn, A.E. (2004) Characterisation of the saponin hydrolysing enzyme avenacoside-α-L-rhamnosidase from the fungal pathogen of cereals, *Stagonospora avenae*. Eur. J. Plant. Pathol. *110*, 421–427.

Jang, I.-S. and Kim, D.-H. (1996) Purification and characterization of α-L-rhamnosidase from *Bacteroides* JY-6, a human intestinal bacterium. Biol. Pharm. Bull. *19*, 1546–1549.

Jansson, P.-E. and Lindberg, B. (1983) Structural studies of gellan gum, an extracellular polysaccharide elaborated by *Pseudomona elodea*. Carbohydr. Res. *124*, 135–139.

Kaji, A. and Ichimi, T. (1973) α-L-Rhamnosidase activity in culture filtrate of *Corticium rolfsii*. Enzymic activity at low pH. Agr. Biol. Chem. *37*, 431–432.

Kim, J.H., Curtis-Long, M.J., Seo, W.D., Lee, J.H., Lee, B.W., Yoon, Y.J., Kang, K.Y. and Park, K.H. (2005) α-L-Rhamnosidase inhibitory activities of polyhydroxylated pyrrolidine. Bioorg. Med. Chem. Lett. *15*, 4282–4285.

Kumar, S., Tamura, K. and Nei, N. (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. *5*, 150–63.

Kurosawa, Y., Ikeda, K. and Egami, F. (1973) α-L-Rhamnosidases of the liver of *Turbo cornutus* and *Aspergillus niger*. J. Biochem. *73*, 31–37.

Manzanares, P., de Graaff, L.H. and Visser, J. (1997) Purification and characterization of an α-L-rhamnosidase from *Aspergillus niger*. FEMS Microbiol. Lett. *157*, 279–283.

Manzanares, P., Orejas, M., Gil, J.V., de Graaff, L.H., Visser, J. and Ramón, D. (2003) Construction of a genetically modified wine yeast strain expressing the *Aspergillus aculeatus rha*A gene, encoding an α-L-rhamnosidase of enological interest. Appl. Environ. Microbiol. *69*, 7558–7562.

Manzanares, P., Orejas, M., Ibáñez, E., Vallés, S. and Ramón, D. (2000) Purification and characterization of an α-L-rhamnosidase from *Aspergillus nidulans*. Lett. Appl. Microbiol. *31*, 198–202.

Manzanares, P., van den Broeck, H., de Graaff, L.H. and Visser, J. (2001) Purification and characterization of two different α-L-rhamnosidases, RhaA and RhaB, from *Aspergillus aculeatus*. Appl. Environ. Microbiol. *67*, 2230–2234.

Martearena, M.R., Blanco, S. and Ellenrieder, G. (2003) Synthesis of alkyl-α-L-rhamnosides by water soluble alcohols enzymatic glycosylation. Bioresource Technol. *90*, 297–303.

Matsumoto, H., Hanamura, S. and Hirayama, M. (2002) Enzymatic production of purified anthocyanins and crystalline products from natural resources, for use in (functional) foods, beverages, drugs and cosmetics. Patent Number WO200222847-A.

McMahon, H., Zoecklein, B.W., Fugelsang, K. and Jasinski, Y. (1999) Quantification of glycosidase activity in selected yeasts and lactic acid bacteria. J. Ind. Microbiol. Biotechnol. *23*, 198–203.

Miake, F., Murata, K., Kuroiwa, A., Kumamoto, T., Kuroda, S., Terasawa, T., Tone, H. and Watanabe, K. (1995) Characterization of *Pseudomonas paucimobilis* FP2001 which forms flagella depending upon the presence of rhamnose in liquid medium. Microbiol. Immunol. *39*, 437–442.

Miake, F., Satho, T., Takesue, H., Yanagida, F., Kashige, N. and Watanabe, K. (2000) Purification and characterization of intracellular α-L-rhamnosidase from *Pseudomonas paucimobilis* FP2001. Arch. Microbiol. *173*, 65–70.

Miklosy, E. and Polos, V. (1995) Yeasts with β-D-glucosidase activity: properties and posible aplication in winemaking processes. Acta Alimentaria *24*, 167–180.

Mimaki, Y., Kuroda, M., Kameyama, A., Yokosuka, A. and Sashida, Y. (1998) New steroidal constituents of the underground parts of *Ruscus aculeatus* and their cytostatic activity on HL-60 cells. Chem. Pharm. Bull. *46*, 298–303.

Miyata, T., Kashige, N., Satho, T., Yamaguchi, T., Aso, Y. and Miake F. (2005) Cloning, sequence analysis, and expression of the gene encoding *Sphingomonas paucimobilis* FP2001 α-L-rhamnosidase. Curr. Microbiol. *51*, 105–109.

Monti, D., Pisvejcová, A., Kren, V., Lama, M. and Riva, S. (2004) Generation of an α-L-rhamnosidase library and its application for the selective derhamnosylation of natural products. Biotechnol. Bioeng. *87*, 763–771.

Morrissey, J.P., Wubben, J.P. and Osbourn, A.E. (2000) *Stagonospora avenae* secretes multiple enzymes that hydrolyse oat leaf saponins. Mol. Plant Microbe Interact. *13*, 1041–1052.

Mutter, M., Beldman, G., Schols, H.A. and Voragen, A.G.J. (1994) Rhamnogalacturonan α-L-rhamnopyranohydrolase. A novel enzyme specific for the terminal nonreducing rhamnosyl unit in rhamnogalacturonan regions of pectin. Plant Physiol. *106*, 241–250.

Norouzian, D., Hosseinzadeh, A., Inanlou, D.N. and Moazami, N. (1999) Various techniques used to immobilize naringinase produced by *Penicillium decumbens* PTCC 5248. World J. Microbiol. Biotechnol. *15*, 501–502.

Ochsner, U.A., Fiechter, A. and Reiser, J. (1994) Isolation, characterization, and expression in *Escherichia coli* of the *Pseudomonas aeruginosa rhlAB* genes encoding a rhamnosyltransferase involved in rhamnolipid biosurfactant synthesis. J. Biol. Chem. *31*, 19787–19795.

Oda, Y., Saito, K., Ohara-Takada, A. and Mori, M. (2002) Hydrolysis of the potato glycoalkaloid α-chaconine by filamentous fungi. J. Biosci. Bioeng. *94*, 321–325.

Orejas, M., Ibáñez, E. and Ramón, D. (1999) The filamentous fungus *Aspergillus nidulans* produces an α-L-rhamnosidase of potential oenological interest. Lett. Appl. Microbiol. *28*, 383–388.

Osbourn, A.E. (1996a) Saponins and plant defense-a soap history. Trends Plant Sci. *1*, 4–9.

Osbourn, A.E. (1996b) Preformed antimicrobial compounds and plant defense against fungal attack. Plant Cell *8*, 1821–1831.

Pearson, W.R. (1990) Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol. *183*, 63–98.

Pisvejcová, A., Monti, D., Sedmera, P., Tarabiono, C., Lama, M., Riva, S. and Kren, V. (2003) α-L-rhamnosidases and their use in selective trimming of natural compounds. 12th European Carbohydrate Symposium. Grenoble, France, July 6–11. PC022.

Prakash, S., Singhal, R.S. and Kulkarni, P.R. (2002) Enzymic debittering of Indian grapefruit (*Citrus paradis*) juice. J. Sci. Food Agric. *82*, 394–397.

Puri, M., Kaur, H. and Kennedy, J.F. (2005) Covalent immobilization of naringinase for the transformation of a flavonoid. J. Chem. Technol. Biotechnol. *80*, 1160–1165.

Puri, M., Marwaha, S.S., Kothari, R.M. and Kennedy, J.F. (1996) Biochemical basis of bitterness in citrus fruit juices and biotech approaches for debbitering. Crit. Rev. Biotechnol. *16*, 145–155.

Qian, S., Yu, H., Zhang, C., Lu, M., Wang, H. and Jin, F. (2005) Purification and characterization of dioscin-α-L-rhamnosidase from pig liver. Chem. Pharm. Bull. *53*, 911–914.

Renault, J.-H., Thepenier, P., Zeches-Hanrot, M., Le Men-Olivier, L., Durand, A., Foucault, A. and Margraft, R. (1997) Preparative separation of anthocyanins by gradient elution centrifugal partition chromatography. J. Chromatogr. A *763*, 345–352.

Ridley, B.L., O'Neill, M.A. and Mohnen, D. (2001) Pectins: structure, biosynthesis and oligogalacturonide-related signalling. Phytochem. *57*, 929–967.

Rodríguez, M.E., Lopes, C.A., van Broock, M., Vallés, S., Ramón, D. and Caballero, A.C. (2004) Screening and typing of patagonian wine yeasts for glycosidase activities. J. Appl. Microbiol. *96*, 84–95.

Romero, C., Manjón, A., Bastida, J. and Iborra, J.L. (1985) A method for assaying the rhamnosidase activity of naringinase. Anal. Biochem. *149*, 566–571.

Sánchez, M.A., Romero, C., Manjón, A. and Iborra, J.L. (1987) Activity of soluble and immobilized hesperidinase on insoluble hesperidin. Biotechnol. Lett. *9*, 871–874.

Scaroni, E., Cuevas, C., Carrillo, L. and Ellenrieder, G. (2002) Hydrolytic properties of crude α-L-rhamnosidases produced by several wild strains of mesophilic fungi. Lett. Appl. Microbiol. *34*, 461–465.

Soares, N.F.F. and Hotchkiss, J.H. (1998) Naringinase immobilization in packaging films for reducing naringin concentration in grapefruit juice. J. Food Sci. *63*, 61–65.

Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R. and Reinemer, P. (1994) Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. Science *265*, 383–386.

Suzuki, H. (1962) Hydrolysis of flavonoid glycosides by enzymes: rhamnodiastase from *Rhamnus* and other sources. Arch. Biochem. Biophys. *99*, 476–483.

Takatsu, T., Takahashi, S., Takamatsu, Y., Shioiri, T., Iwado, S. and Haneishi, T. (1987a) Chloropolysporins A, B and C, novel glycopeptide antibiotics from *Faenia interjecta* p.nov. IV. Partially deglycosylated derivatives. J. Antibiot. (Tokyo) *40*, 941–945.

Takatsu, T., Katayama, T., Nakajima, M., Takahashi, S., Haneishi, T., Magaribuchi, T. and Tajima, M. (1987b) Chloropolysporins A, B and C, novel glycopeptide antibiotics from *Faenia interjecta* p.nov. V. Comparative studies of the biological properties. J. Antibiot. (Tokyo) *40*, 946–952.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673–4680.

Trummler, K., Effenberger, F. and Syldatk, C. (2003) An integrated microbial/enzymatic process for production of rhamnolipids and L-(+)-rhamnose from rapeseed oil with *Pseudomonas* sp. DSM 2847. Eur. J. Lipid Sci. Technol. *105*, 563–571.

de Vries, R.P., van Grieken, C., vanKuyk, P.A. and Wösten, H.A. B. (2005) The value of genome sequences in the rapid identification of novel genes encoding specific plant cell wall degrading enzymes. Curr. Genom. *6*, 157–187.

Yanai, T. and Sato, M. (2000) Purification and characterization of an α-L-rhamnosidase from *Pichia angusta* X349. Biosci. Biotechnol. Biochem. *64*, 2179–2185.

Yokozawa, T. and Owada, S. (1999) Effects of ginsenoside Rd in cephaloride-induced renal disorder. Nephron *81*, 200–207.

Young, N.M., Johnston, R.A.Z. and Richards, J.C. (1989) Purification of the α-L-rhamnosidase of *Penicillium decumbens* and characteristics of two glycopeptide components. Carbohydr. Res. *191*, 53–62.

Yu, H., Gong, J., Zhang, C. and Jin, F. (2002) Purification and characterization of ginsenoside-α-L-rhamnosidase. Chem. Pharm. Bull. *50*, 175–178.

Yu, H., Liu, H., Zhang, C., Tan, D., Lu, M. and Jin, F. (2004) Purification and characterization of gypenoside-α-L-rhamnosidase hydrolysing gypenoside-5 into ginsenoside Rd. Process Biochem. *39*, 861–867.

Zverlov, V.V., Hertel, C., Bronnenmeier, K., Hroch, A., Kellermann, J. and Schwarz, W.H. (2000) The thermostable α-L-rhamnosidase RamA of *Clostridium stercorarium*: biochemical characterization and primary structure of a bacterial α-L-rhamnoside hydrolase, a new type of inverting glycoside hydrolase. Mol. Microbiol. *35*, 173–179.

CHAPTER 9

APPLICATION OF GLYCOSIDASES AND TRANSGLYCOSIDASES IN THE SYNTHESIS OF OLIGOSACCHARIDES

FRANCISCO J. PLOU, ARÁNZAZU GÓMEZ DE SEGURA AND ANTONIO BALLESTEROS*

Departamento de Biocatálisis, Instituto de Catálisis y Petroleoquímica, CSIC, Cantoblanco, Madrid, Spain

**a.ballesteros@icp.csic.es*

1. INTRODUCTION

It is generally accepted that an oligosaccharide is a carbohydrate consisting of 2–10 monosaccharide residues linked by *O*-glycosidic bonds (Eggleston and Cote, 2003; McNaught, 1997). The development of efficient and scalable processes for the synthesis of oligosaccharides is of considerable interest to the food and pharmaceutical industries (Kren and Thiem, 1997; Macmillan and Daines, 2003).

Oligosaccharides are quite complex molecules: 3 different hexopyranose moieties yield up to 720 trisaccharides. For the preparation of structurally well-defined oligosaccharides, the stereo- and regio-specificities of enzymes are very attractive properties compared to chemical processes that require complex protection and deprotection steps. At present, enzymatic processes are preferred in industry for the production of most commercial oligosaccharides.

The synthesis of glycosidic bonds *in vivo* is performed by glycosyltransferases (EC 2.4.) (Ichikawa *et al*., 1992). These enzymes catalyse the transfer of a glycosyl donor to an acceptor molecule forming a new glycosidic bond regio- and stereospecifically. According to the nature of the sugar residue being transferred, glycosyltransferases are divided into hexosyltransferases (EC 2.4.1.), pentosyltransferases (EC 2.4.2.), and those transferring other glycosyl groups (EC 2.4.99.). Depending on the nature of the donor molecule, glycosyltransferases are classified into three main mechanistic groups: (1) Leloir-type glycosyltransferases, which require sugar nucleotides (*e.g.* UDP-glucosyltransferases); (2) non-Leloir glycosyltransferases, which use sugar-1-phosphates (*e.g.* phosphorylases); and (3) transglycosidases,

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 141–157

which employ non-activated sugars such as sucrose, lactose or starch. A distinctive feature of transglycosidases, compared with Leloir and non-Leloir glycosyltransferases, is that they also display some hydrolytic activity that can be regarded as a transfer of a glycosyl group from the donor to water.

In addition to glycosyltransferases, under appropriate conditions glycosidases (glycoside hydrolases, EC 3.2) can also be used for *in vitro* synthesis of oligosaccharides. Glycosidases catalyse the hydrolysis of glycosidic bonds in oligo- and polysaccharides with a high degree of stereospecificity. It is noteworthy that, in terms of reaction mechanism, transglycosidases and glycosidases belong to the same group. In fact, transglycosidases and glycosidases are grouped in the 'glycoside hydrolase family (GH family)' in the Henrissat classification, which is based on amino acid sequence comparisons (Coutinho and Henrissat, 1999). The GH family comprises more than 2500 enzymes.

Several reviews on oligosaccharide synthesis by glycosyltransferases of the Leloir and non-Leloir types have been published recently (Daines *et al*, 2004; Hamilton, 2004). The main problems associated with such syntheses are (1) the requirement for sugar nucleotides or sugar phosphates as substrates, (2) the inhibitory effect of the nucleotide phosphate released, and (3) the limited availability of these enzymes. Nevertheless, continuous progress in the study of these enzymes, the cloning of new variants, and the application of molecular evolution and site-directed mutagenesis for better performance is improving their potential for use in oligosaccharide synthesis (Planas and Faijes, 2002). In this chapter we will focus on the use of transglycosidases and glycosidases for the synthesis of oligosaccharides because they both employ the same type of substrate and share the same mechanism of glycosylation (Sanz-Aparicio *et al*, 1998).

2. SYNTHESIS OF OLIGOSACCHARIDES BY GLYCOSIDASES

Glycosidases are classified as being either retaining (*e.g.* β-galactosidase) or inverting (*e.g.* trehalase) because enzymatic hydrolysis of glycoside bonds can proceed with the net retention or net inversion of the anomeric configuration, respectively. Glycosidases are widely employed for oligosaccharide synthesis because, under appropriate conditions, the normal hydrolytic reaction can be reversed towards glycosidic bond synthesis (Ajisaka and Yamamoto, 2002; Scigelova *et al*, 1999). Most glycosidases used for synthetic purposes are retaining glycosidases, and in particular exo-glycosidases. The glycosyl donor can be a monosaccharide, an oligosaccharide or an activated glycoside (Thiem, 1995). Glycosidase-catalysed oligosaccharide synthesis can be controlled by thermodynamic and kinetic factors as explained below.

2.1. Thermodynamic Synthesis (Reverse Hydrolysis)

In thermodynamically controlled synthesis (*i.e.* the reaction takes place until it reaches equilibrium), the reaction can proceed between two monosaccharides

(forming a disaccharide) or between one monosaccharide and an alcohol (yielding a glycoside). Water is the leaving group. The hydrolysis/synthesis equilibrium is balanced by approximately 4 kcal/mol towards bond cleavage (Planas and Faijes, 2002). The factors that determine yield in thermodynamically controlled processes are initial substrate concentration, pH, temperature, ionic strength, solvent composition, *etc*.

To shift the equilibrium towards synthesis, one (or various) of the following strategies has been followed: a) the use of high substrate concentrations; b) the addition of an organic co-solvent to reduce water activity (a small amount of water is nevertheless required to maintain enzyme activity and to dissolve the carbohydrates); c) use of the acceptor (an alcohol) as the reaction medium; d) the use of high temperatures – a number of thermostable glycosidases have been characterized in recent years, in particular the β-glucosidase from the hyperthermophilic archeon *Pyrococcus furiosus* (Bhatia *et al*., 2002)–. However, most of these approaches are compromised by loss of enzyme activity/stability and reduced sugar solubility.

Reverse hydrolysis is economically feasible and simple because the enzymes required are readily available and inexpensive (*e.g.* β-galactosidases are cheap enzymes industrially used in the hydrolysis of lactose, and also applicable to galactooligosaccharide synthesis). However, the yields obtained are usually low ($\leq$ 20%). Crout and collaborators synthesized a variety of β-D-glucosides (maximum yield 20%) using almond β-D-glucosidase in a medium containing 80–90% (v/v) organic solvent (acetonitrile or *tert*-butanol) (Vic *et al*., 1996). Using the acceptor alcohol as solvent, the thermodynamic equilibrium was shifted towards synthesis (40–60% yield) not only by mass action but also by the reduced water activity (Vic and Crout, 1995).

Interestingly, the synthetic specificity of many glycosidases may differ substantially from the specificity of hydrolysis (Ajisaka and Yamamoto, 2002). Thus, many α-glucosidases – the function of which in nature is the hydrolysis of $\alpha(1 \rightarrow 4)$ bonds– are able to transfer glucose units to the primary (more reactive) 6-OH of the acceptor, yielding products such as isomaltose, panose, *etc*. In addition, transfer to secondary hydroxyl groups (2-OH, 3-OH, 4-OH) usually takes place and, as a result, a mixture of oligosaccharides consisting of $\alpha(1 \rightarrow 2)$, $\alpha(1 \rightarrow 3)$ and $\alpha(1 \rightarrow 4)$ bonds is obtained (Kato *et al*., 2002). Fig. 1 shows the products formed upon condensation of glucose catalysed by the α-glucosidase from *B. stearothermophilus*. The main product (51%) of the reverse reaction is isomaltose which has an $\alpha(1 \rightarrow 6)$ linkage (Mala and Kralova, 2000).

It has been noted that an enzyme having a poor ability to hydrolyse a tetrasaccharide is unlikely to be able to synthesize such molecules, as the binding conditions for the enzyme-substrate complex will be the same in both reactions (Mala *et al.*, 1999). Most of the examples reported using the thermodynamic approach concern the preparation of disaccharides or glycosides of simple hydrophilic alcohols using exo-glycosidases (β-galactosidases, α- and β-glucosidases and α-mannosidases).

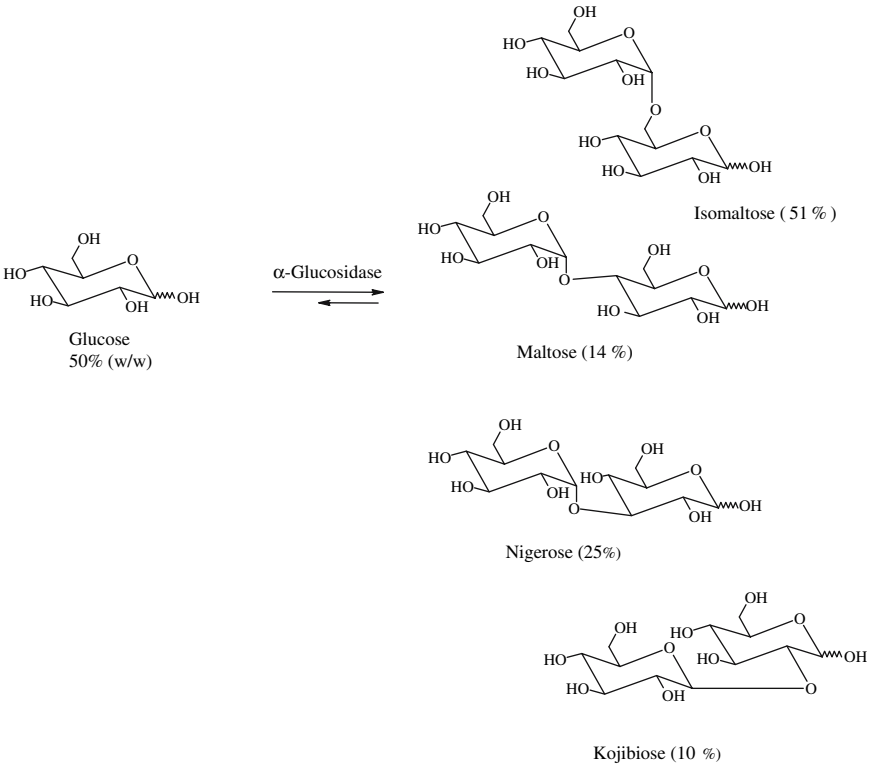


Figure 1. Gluco-oligosaccharide synthesis by reverse-hydrolysis catalysed by α-glucosidase from *Bacillus stearothermophilus.* Conditions: 50% (w/w) glucose solution in 0.1 M phosphate buffer pH 7.5, 10 days. Data derived from Mala and Kralova (2000)

2.2. Kinetic Synthesis (Transglycosylation)

Although reverse hydrolysis has the advantage of simplicity, greater versatility can be obtained using activated glycosides as glycosyl donors. The transglycosylation mechanism of retaining glycosidases, also valid for transglycosidases, is represented in Fig. 2. As a consequence of this mechanism, the anomeric configuration of the resulting oligosaccharide is identical to that of the original donor. The partitioning of the glycosyl-enzyme intermediate between hydrolysis and transfer products is determined by the ratio $k_2 \cdot [H_2O]/k_3 \cdot$ [Acceptor], as can be inferred from Fig. 2. The ratio transferase/hydrolase thus depends on two parameters: the concentration of the acceptor and properties intrinsic to the enzyme *i.e.* its ability to bind the sugar acceptor and to exclude H_2O.

As the reactant is consumed the concentration of the product increases until it reaches a maximum. At this point, the rate of synthesis of the product (k_3) equals its rate of hydrolysis (k_{-3}). Subsequently, kinetic control is lost and the reaction must be stopped quickly before product hydrolysis becomes the major process.

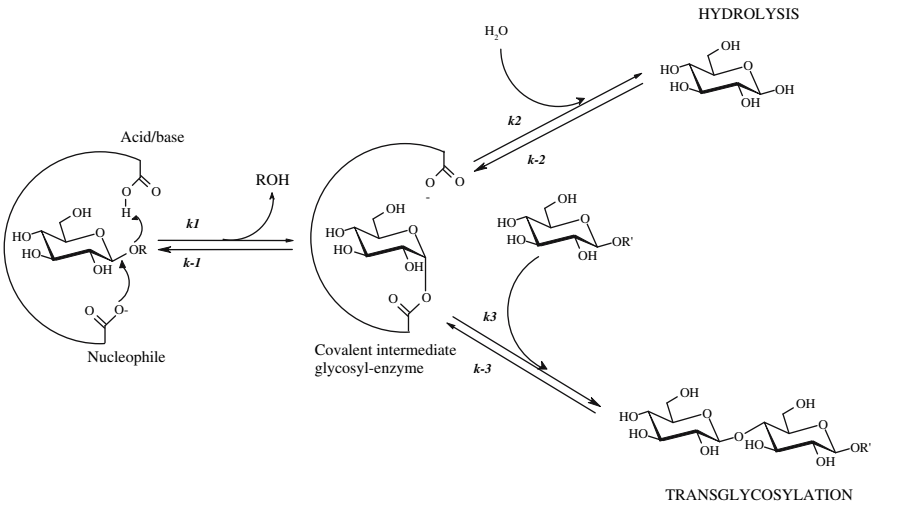


Figure 2. Mechanism of transglycosylation catalysed by retaining glycosidases and transglycosidases. The active site contains two carboxylic acid residues, located approximately 5.5 Å apart: one acting as a nucleophile and the other as an acid/base catalyst. The reaction proceeds by a double-displacement mechanism in which a covalent glycosyl-enzyme intermediate is formed by the attack of the deprotonated carboxylate on the anomeric centre of the carbohydrate with concomitant breaking of the scissile C-O glycosidic bond. This step is assisted by the carboxylic residue acting as general acid. The second step is the attack of a nucleophile (a carbohydrate) on the glycosyl-enzyme intermediate, which is assisted by the conjugate base of the second carboxyl residue

The existence of this maximum explains why transglycosylation results in higher yields of condensation products compared with equilibrium-controlled processes. Synthetic yields using kinetic approaches are usually close to 40% compared with the 20% yield typically obtained in thermodynamically-controlled processes. However, reaction time must be carefully controlled as hydrolysis subsequently predominates (Bruins *et al*., 2004).

To reduce the extent of hydrolysis, several approaches can be attempted: (1) continuous removal of the transglycosylation product by crystallization, selective adsorption onto different carriers or coupling to another enzymatic process (Planas and Faijes, 2002); (2) the presence of a suitable glycosyl acceptor that reacts as a nucleophile faster than water; (3) the use of high concentrations of acceptor (Cobucci-Ponzano *et al*., 2003). Another common approach is the use of activated donors which are rapidly and irreversibly cleaved so that $k_{-1} \approx 0$ (Fig. 2). Examples include *o*- and *p*-nitrophenyl glycosides, vinyl glycosides, glycosyl fluorides or glycals (Boons and Isles, 1996; Shoda *et al*., 2003). These substrates have the advantage that the leaving group (fluoride, phenol) is a poor acceptor and will not compete with the actual acceptor molecule. In addition, the activated sugar is a much better substrate than the product formed. However, some glycosidases do not accept activated substrates but only disaccharide or higher oligosaccharide glycosyl donors.

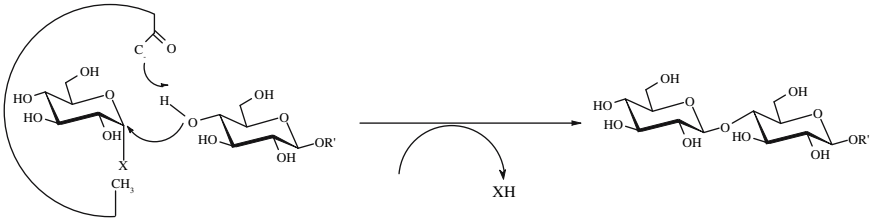


Figure 3. Mechanism of transglycosylation catalysed by glycosynthases. The donor sugar is an activated glycosyl donor with an anomeric configuration opposite to that of the normal substrate (*i.e.* an α-glycosyl fluoride for a β-glycosynthase), thus mimicking the covalent intermediate glycosyl-enzyme. This is followed by the attack of a nucleophile on the glycosyl fluoride, yielding a disaccharide. The reaction is irreversible because the product formed cannot react with the active site as the catalytic nucleophile is not present in the glycosynthase

2.3. Glycosynthases: A Special Case

As a consequence of the progress made in understanding of the structures and catalytic mechanisms involved in the enzymatic synthesis of glycosidic bonds, a group of novel, site-specifically mutated glycosidases called glycosynthases were developed (Davies *et al*, 2001). The glycosynthase concept was introduced in 1998 by Withers and collaborators using an exo-glucosidase (Mackenzie *et al*, 1998) and extended shortly thereafter to endo-glycosidases (Malet and Planas, 1998).

A glycosynthase is a specifically-mutated retaining glycosidase in which substitution of the catalytic carboxyl nucleophile by a non-nucleophilic residue (Ala, Gly or Ser) results in an enzyme which is hydrolytically inactive but yet able to catalyse the transglycosylation of activated glycosyl fluoride donors having the opposite anomeric configuration to that of the normal glycosidase substrate. To convert a glycosidase into a glycosynthase, it is thus necessary to identify the residue acting as the catalytic nucleophile. The enzyme-substrate complex in glycosynthases mimics the glycosyl-enzyme intermediate formed by retaining glycosidases and is able to react with an acceptor (normally a carbohydrate) in a similar way to the transglycosylation step performed by the retaining glycosidases (Fig. 3). By this means the desired oligosaccharide accumulates and yields obtained can reach 95–98% in some cases (Planas and Faijes, 2002). The impressive number of glycosidases available clearly indicates that the potential biodiversity of glycosynthases is very largely unexplored, and novel applications of these enzymes will undoubtedly emerge (Perugino *et al*, 2005). Very recently, the first glycosynthase derived from an inverting glycosidase has been reported (Honda and Kitaoka, 2006).

3. SYNTHESIS OF OLIGOSACCHARIDES BY TRANSGLYCOSIDASES

Transglycosidases are ideal biocatalysts for oligosaccharide synthesis *in vitro* since they do not require specially activated substrates but directly employ the free energy of cleavage of disaccharides (*e.g.* sucrose) or polysaccharides (*e.g.* starch)

(Plou *et al*., 2002). Transglycosidases present the same mechanism as retaining glycosidases (Fig. 2), resulting in the net retention of the anomeric configuration. Although the normal function of transglycosidases is the transfer of glycosyl residues, water may also act as the acceptor of the glycosyl-enzyme intermediate. In fact the assignation of oligosaccharide-producing enzymes as either glycosidases or transglycosidases still remains controversial. For a particular enzyme to be designed a transglycosidase it must possess a significant ability to bind the acceptor and exclude H_2O. The most important groups of transglycosidases are transglucosidases and transfructosidases.

3.1. Transglucosidases

Glucansucrases (EC 2.4.1.5) and cyclodextrin glucosyltransferase (CGTase, EC 2.4.1.19) are the most representative enzymes of the transglucosidase family, the natural substrates of which are sucrose and starch, respectively.

3.1.1. Glucansucrases

Several bacteria excrete a range of transglucosidases called glucansucrases that utilise sucrose as the sole energy source to synthesise glucose polymers. Glucansucrases belong to family 70 of the glycoside hydrolase family in the Henrissat classification. Glucansucrases from streptococci are involved in the formation of dental caries (Devulapalle *et al*., 2004). Dextransucrases (sucrose:1,6-α-D-glucan 6-α-D-glucosyltransferase) are glucansucrases produced by different *Leuconostoc mesenteroides* strains that convert sucrose into α(1 → 6)-linked glucose polymers (dextrans), releasing fructose (Monchois *et al*., 1999). However, other short carbohydrates may also act as acceptors yielding the so-called acceptor products (Robyt and Walseth, 1978). The three reactions catalysed by dextransucrase, (a) polymerisation of the glucose moiety of sucrose, (b) glucose transfer to acceptors, and (c) sucrose hydrolysis, are competitive. Some acceptors (*e.g.* isomaltose) yield a homologous series of oligosaccharides, presenting an increasing number of glucose moieties in their structure; others form a unique acceptor-product containing one glucose residue more than the acceptor (Robyt, 1996). The latter is the case for fructose, which is a major product in all dextransucrase-catalysed reactions. Fructose yields leucrose (α-D-Glu*p*-(1 → 5)-D-Fru*p*) along with a minor product, isomaltulose (α-D-Glu*p*-(1 → 6)-D-Fru*p*). The leucrose synthesis process becomes particularly important in the final stages of dextransucrase-catalysed syntheses because the fructose concentration is high (Buchholz *et al*., 1998). Acceptors are classified as being strong (*e.g.* maltose), which enhance the reaction rate (measured as fructose released) and strongly inhibit the synthesis of dextran, or weak (*e.g.* fructose), which have an inhibitory effect on glucan formation but yield small amounts of acceptor-products (Monchois *et al*., 1999).

The regioselectivity displayed by dextransucrases is highly strain dependent (Jeanes *et al*., 1954). The dextransucrase from *L. mesenteroides* NRRL B-512F synthesises α(1 → 6) linked gluco-oligosaccharides (Robyt and Eklund, 1983).

With several acceptors such as glucose, methyl 1-O-α-D-glucopyranoside, maltose or isomaltose a series of isomaltodextrins with a degree of polymerisation ranging from 2 to 7 is obtained. Isomalto-oligosaccharides constitute an important group of oligosaccharides used as prebiotics, immunostimulants and anti-caries agents (Goulas *et al*., 2004). Dextransucrase from strain B-1299 is also able to form α(1 → 2) linkages (Dols-Lafargue *et al*., 2001; Gómez de Segura *et al*., 2003). which confer particular properties (Boucher *et al*., 2003; Djouzi *et al*., 1995; Simmering and Blaut, 2001). Gluco-oligosaccharides containing α(1 → 2) bonds are capable of promoting the selective development of beneficial cutaneous flora. Based on the acceptor reaction with maltose, dextransucrase from *L. mesenteroides* B-1299 is being exploited to produce 50 Tm/year of non-digestible gluco-oligosaccharides containing α(1 → 2) bonds for the dermo-cosmetic industry (Dols *et al*., 1998).

3.1.2. CGTases

Cyclodextrin glucanotransferases (CGTases) constitute a group of transglucosidases that belong to family 13 of the glycoside hydrolases (α-amylase family): This family includes different starch-processing enzymes comprising α-amylases, α-glucosidases, pullulanases and isoamylases. All members of family 13 contain a $(\beta/\alpha)_8$-barrel catalytic domain (Leemhuis and Dijkhuizen, 2003). CGTases catalyse the formation of cyclodextrins (CDs) from starch by an intramolecular transglucosylation reaction (*cyclization*) in which part of the α(1 → 4)-amylose chain is cyclized as a result of the formation of an additional α(1 → 4)-glucosidic bond. CDs are excellent encapsulating agents and are widely used in the food, pharmaceutical, chemical and cosmetic industries. CGTases usually produce a mixture of α, β, and γ-CDs (containing six, seven and eight α-D-glucose units respectively). For example, the CGTase from *Thermoanaerobacter* sp. (commercialised as Toruzyme by Novozymes A/S) converts a 25% (w/v) starch dispersion into a mixture of α, β and γ-CDs with an overall yield of 30%.

Apart from the cyclization process, CGTases also catalyse intermolecular transglucosylations using a cyclodextrin (*coupling* reaction) or a linear maltooligosaccharide (*disproportionation* reaction) as glucosyl donors (van der Veen *et al*., 2000). In addition, CGTases catalyse the hydrolysis of starch and maltooligosaccharides, although at a much lower rate (Alcalde *et al*., 1998). Fig. 4 represents the specific activity of a CGTase from *Thermoanaerobacter* sp. in the above reactions. As shown, the greatest activity (approx. 1200 U/mg) corresponds to the transglucosylation between two maltooligosaccharides.

Elucidation of the three-dimensional structure and the biochemical characterization of site-specific mutants have provided detailed insight into the mechanisms of the reactions catalysed by CGTases (Leemhuis and Dijkhuizen, 2003). A distinctive feature of CGTases is the existence of the so-called cyclization axis (generally an aromatic residue, either Phe or Tyr) which is crucial for cyclodextrin formation. Two carboxylic residues (the catalytic nucleophile Asp229 and the acid/base catalyst Glu257) are involved in a combined attack on a glycosidic bond that results in the release of the reducing end of amylose.

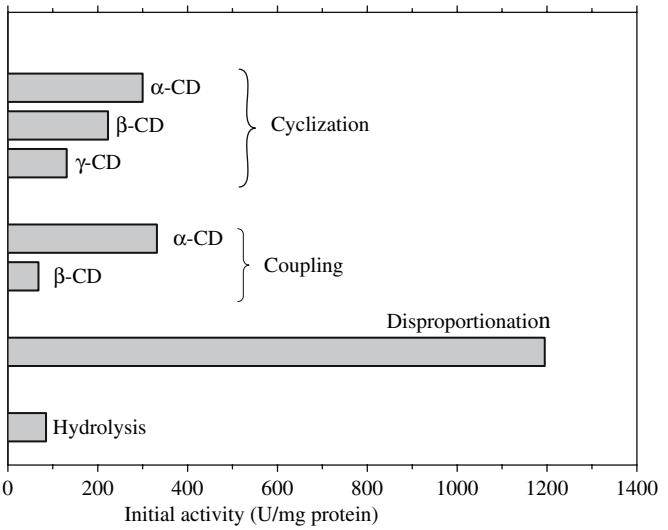


Figure 4. Specific activity of CGTase from *Thermoanaerobacter* sp. Cyclization: formation of α-, β- and γ-CD from starch. Coupling: transglucosylation of α- or β-CD to methyl-α-D- glucopyranoside. Disproportionation: transglycosylation of p-nitrophenyl-α-D-maltoheptaoside-4,6-*O*-ethylidene to maltose. Hydrolysis: potato soluble starch as substrate (average degree of polymerization 50). Data derived from Alcalde *et al* (1999)

A feasible explanation for the differences observed in the transferase/hydrolase ratio within the α-amylase family is the variation in the accessibility of the active site to water (Leemhuis and Dijkhuizen, 2003). This may be related to the hydrophobicity of the residues in the vicinity of the catalytic site and, in particular, near the acid/base catalyst, as mutations in these residues changed the transferase/hydrolase ratio in a neopullulanase. Recently it has been hypothesized that the separation between Glu257 (the acid/base catalytic residue) and Asp328 (a fully conserved residue that stabilizes the transition state) may determine the hydrolysis/transglycosylation specificities of the α-amylase family (Roujeinikova *et al.*, 2001). This distance is larger in strict transglycosylation enzymes. A third explanation is that the glycosyl-enzyme intermediate is favourably stabilized in transferases, which is not necessary in hydrolases, and that a conformational change in the protein induced by a sugar acceptor is required in the transglycosylation step (Fig. 2) (Leemhuis and Dijkhuizen, 2003). In this context, CGTase has been transformed into a starch hydrolase by directed evolution (Leemhuis *et al*, 2003). Chemical modification of certain CGTase residues has also resulted in increased transglycosylation (Alcalde *et al*, 2001) or hydrolysis (Alcalde *et al*, 1999) activities.

When an acceptor is present in the reaction mixture it inhibits the formation of cyclodextrins. The acceptor specificity of CGTase is rather broad. A number of hydroxyl-containing compounds such as glycosides, sugar alcohols, vitamins, flavonoids, *etc*. may act as CGTase acceptors, in many cases with high efficiency (Aga

et al., 1991; Kim *et al.*, 1997). The transglucosylation activity of CGTase seems to be very dependent on enzyme source (Park *et al.*, 1998). Glucosylation often confers new stability/solubility properties to an aglycon (Kometani *et al.*, 1994). However, the best acceptors are carbohydrates with an α-D-glucopyranose structure in the chair form and equatorial hydroxyl groups at C-2, C-3 and C-4 (Tonkova, 1998). With maltose or glucose as acceptors and starch as donor, a series of maltooligosaccharides is produced (Martin *et al.*, 2001). The degree of polymerisation of the oligosaccharides formed can be modulated by varying the starch to acceptor ratio. CGTase has a higher affinity for disaccharides compared to monosaccharides which suggests that the acceptor binding site can accommodate at least two glucopyranose moieties (Park *et al.*, 1998). For example, disaccharides such as isomaltose, gentiobiose, turanose, maltulose, isomaltulose, cellobiose and sucrose are good CGTase acceptors. A steric factor possibly plays a major role in diminishing the acceptor capacity of trisaccharides.

3.2. Transfructosidases

Many micro-organisms and approx. 12% ($4 \cdot 10^4$ species) of higher plants build carbohydrate stores based on fructans which are formed by β-D-fructofuranose units with a terminal D-glucose. The fructosyl moieties can be β(2 → 6)-linked as is the case for levan, or β(2 → 1)-linked as in inulin. These compounds are synthesized by transfructosidases called levansucrases and inulosucrases, respectively (Olivares-Illana *et al.*, 2002; Tungland, 2003). Both enzymes utilize sucrose as the energy source for fructan synthesis.

In addition, a group of transfructosidases that are produced by fungi (*Aureobasidium pullulans, Aspergillus niger, Aspergillus oryzae*, *etc.*) catalyse the synthesis of short-chain fructo-oligosaccharides (FOS) (Fernandez *et al.*, 2004; Sangeetha *et al.*, 2005; Shin *et al.*, 2004). FOS of the inulin-type are fructose oligomers with a terminal glucose group in which 2–4 fructosyl moieties are linked via β(2 → 1)-glycosidic bonds (Antosova and Polakovic, 2001). Commercial FOS are mainly composed of 1-kestose (GF_2), nystose (GF_3) and 1^F-fructofuranosyl-nystose (GF_4). FOS are non-cariogenic and have a sweetness about 40–60% that of sucrose. They are produced at multi-ton scale given their use as prebiotics. Prebiotic agents are food ingredients that escape hydrolysis in the upper gastrointestinal tract, enter the colon, and produce positive effects on human health because they are selectively fermented by beneficial colonic flora (*Bifidobacterium* and *Lactobacillus*). As a consequence of the metabolism of these bacteria, short-chain fatty acids (acetate, propionate and butyrate) and L-lactate are produced (Probert and Gibson, 2002) with the following implications for health (Gibson and Ottaway, 2000; Tuohy *et al.*, 2005): (1) potential protective effects against colorectal cancer and bowel infectious diseases by inhibiting putrefactive and pathogen bacteria; (2) improvement of the bioavailability of essential minerals; (3) enhancement of glucid and lipid metabolism.

FOS-producing enzymes belong to families 32 and 68 of the glycoside hydrolases. Assignation of FOS-producing enzymes as β-fructofuranosidases (EC 3.2.1.26) or

transfructosidases –fructosyltransferases- (EC 2.4.1.9) still remains in dispute. The assignation of a particular enzyme as a β-fructofuranosidase or a transfructosidase should be based on the transferase to hydrolysis ratio, especially at low substrate concentrations. In fact, only a few of these enzymes have a transfructosylating activity significant enough for industrial FOS production. Recently, several FOS-synthesizing enzymes from *Aspergillus* species have been purified and characterized (Velasco and Adrio, 2002), and the first three-dimensional structure of a β-fructofuranosidase, namely that of *Thermotoga maritima*, has been reported (Alberto *et al*., 2004).

Maximal FOS production for any particular enzyme depends on the relative rates of the transfructosylation and hydrolysis reactions (Nguyen *et al*., 2005). Ghazi *et al.* (2005), using an immobilized transfructosidase and 630 g/l sucrose, obtained a maximum FOS production of 61.5% (w/w), referred to the total amount of carbohydrates in the mixture. At the point of maximum FOS concentration, the weight ratio 1-kestose/nystose/1^F-fructofuranosylnystose was 6.2/3.7/0.1. Similar yields of fructo-oligosaccharides have been reported with other immobilized transfructosidases (Chiang *et al*., 1997; Tanriseven and Aslan, 2005).

Levansucrases catalyse the synthesis of levan from sucrose, a polymer with applications in medicine, pharmacy, agriculture and food (Steinbchel and Rhee, 2005). In addition to levan formation, levansucrases concomitantly produce FOS of the inulin-type (Euzenat *et al*., 1997; Tambara *et al*., 1999; Trujillo *et al*., 2001), and also catalyse other transfructosylation reactions in the presence of acceptors such as methanol (Kim *et al*., 2000), glycerol (Gonzalez-Munoz *et al*., 1999) and disaccharides (Park *et al*., 2003). Levansucrases are included in glycoside hydrolase (GH) family 68. The crystal structure of *Bacillus subtilis* levansucrase was recently solved by Meng and Fütterer (2003) at 1.5 Å resolution, and shows a rare five-bladed β-propeller. Site-directed mutations of the three putative catalytic residues of the *Lactobacillus reuteri* 121 levansucrase and inulosucrase (the catalytic nucleophile, the general acid/base catalyst, and the transition state stabilizer) have been obtained recently (Ozimek *et al*., 2004).

Neo-fructo-oligosaccharides (neo-FOS) consist mainly of neokestose (neo-GF2) and neonystose (neo-GF3), in which a fructosyl unit is β(2 → 6) bound to the glucose moiety of sucrose or 1-kestose, respectively (Fig. 5). Grizard and Barthomeuf (1999) were the first to report the enzymatic synthesis of neo-FOS using a transfructosylating activity present in a commercial enzyme preparation from *Aspergillus awamori*. The neo-FOS yield reached a maximum of 50% (w/w) based on total weight of carbohydrates in the reaction mixture. Cultures of the astaxanthin-producing yeast *Xanthophyllomyces dendrorhous* accumulated neokestose as a major transfructosylation product when growing on sucrose (Kilian *et al.*, 1996; Kritzinger *et al*., 2003). Neokestose also occurs as a minor transfructosylation product of whole cells or enzymes from various plants, yeasts (*e.g. S. cerevisiae)* and some filamentous fungi (Hayashi *et al*., 2000). Investigation using human faeces as an inoculum *in vitro* have demonstrated that neokestose has prebiotic effects that surpass those of commercial FOS (Kilian *et al*., 2002).

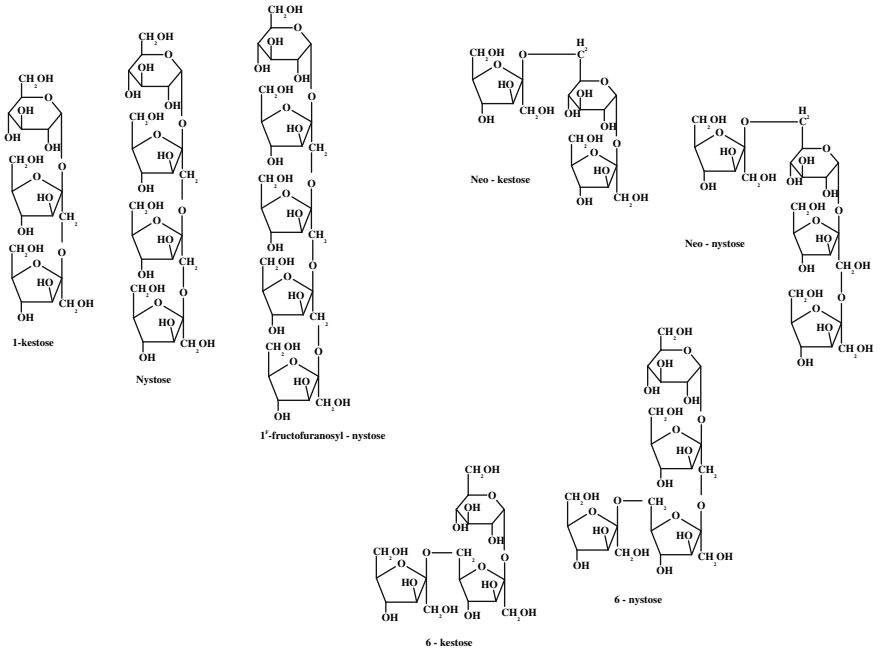


Figure 5. Structure of the inulin-type fructo-oligosaccharides, neo-FOS and 6^F-type FOS

Short-chain ^{6}F-type fructo-oligosaccharides have also received some attention (Fig. 5). Both linear and branched β-(2,6)-linked FOS (the first is 6-kestose) occur naturally in various food products (Marx *et al*., 2000). However, the enzymatic synthesis of ^{6}F-type FOS has barely been studied. Straathof *et al*. (1986) were the first reporting that the invertase from *Saccharomyces cerevisiae* formed 6-kestose at high sucrose concentrations (2.34 M, 800 g/l). Bekers *et al*. (2002) determined the presence of the trisaccharides 1-kestose, neokestose and 6-kestose in the fructan syrup obtained with a levansucrase from the ethanol-producing bacteria *Zymomonas mobilis*.

ACKNOWLEDGEMENTS

This work was supported by the Spanish CICYT (Project BIO2004-03773-C04-01).

REFERENCES

Aga, H., Yoneyama, M., Sakai, S. and Yamamoto, I. (1991) Synthesis of 2-O-alpha-D-glucopyranosyl L-ascorbic-acid by cyclomaltodextrin glucanotransferase from *Bacillus stearothermophilus*. Agric. Biol. Chem. *55*, 1751–1756.

Ajisaka, K. and Yamamoto, Y. (2002) Control of the regioselectivity in the enzymatic syntheses of oligosaccharides using glycosidases. Trends Glycosci. Glycotechnol. *14*, 1–11.

Alberto, F., Bignon, C., Sulzenbacher, G., Henrissat, B. and Czjzek, M. (2004) The three-dimensional structure of invertase (beta-fructosidase) from *Thermotoga maritima* reveals a bimodular arrangement and an evolutionary relationship between retaining and inverting glycosidases. J. Biol. Chem. *279*, 18903–18910.

Alcalde, M., Plou, F.J., Andersen, C., Martin, M.T., Pedersen, S. and Ballesteros, A. (1999) Chemical modification of lysine side chains of cyclodextrin glycosyltransferase from *Thermoanaerobacter* causes a shift from cyclodextrin glycosyltransferase to alpha-amylase specificity. FEBS Lett. *445*, 333–337.

Alcalde, M., Plou, F.J., Martin, M.T., Valdes, I., Mendez, E. and Ballesteros, A. (2001) Succinylation of cyclodextrin glycosyltransferase from *Thermoanaerobacter* sp. 501 enhances its transferase activity using starch as donor. J. Biotechnol. *86*, 71–80.

Alcalde, M., Plou, F.J., Pastor, E. and Ballesteros, A. (1998) Effect of chemical modification of cyclodextrin glycosyltransferase (CGTase) from *Thermoanaerobacter* sp on its activity and product selectivity. Ann. NY Acad. Sci. *864*, 183–187.

Antosova, M. and Polakovic, M. (2001) Fructosyltransferases: The enzymes catalyzing production of fructooligosaccharides. Chemical Papers-Chemicke Zvesti *55*, 350–358.

Bekers, M., Laukevics, J., Upite, D., Kaminska, E., Vigants, A., Viesturs, U., Pankova, L. and Danilevics, A. (2002) Fructooligosaccharide and levan producing activity of *Zymomonas mobilis* extracellular levansucrase. Process Biochem. *38*, 701–706.

Bhatia, Y., Mishra, S. and Bisaria, V.S. (2002) Microbial beta-glucosidases: Cloning, properties, and applications. Crit. Rev. Biotechnol. *22*, 375–407.

Boons, G.J. and Isles, S. (1996) Vinyl glycosides in oligosaccharide synthesis. 2. The use of allyl and vinyl glycosides in oligosaccharide synthesis. J. Org. Chem. *61*, 4262–4271.

Boucher, J., Daviaud, D., Simeon-Remaud, M., Carpene, C., Saulnier-Blache, J.S., Monsan, P. and Valet, P. (2003) Effect of non-digestible gluco-oligosaccharides on glucose sensitivity in high fat diet fed mice. J. Physiol. Biochem. *59*, 169–173.

Bruins, M.E., Strubel, M., van Lieshout, J.F.T., Janssen, A.E.M. and Boom, R.M. (2004) Oligosaccharide synthesis by the hyperthermostable β-glucosidase from Pyrococcus furiosus: kinetics and modeling. Enz. Microb. Technol. *33*, 3–11.

Buchholz, K., Noll-Borchers, M. and Schwengers, D. (1998) Production of leucrose by dextransucrase. Starch-Starke *50*, 164–172.

Chiang, C.J., Lee, W.C., Sheu, D.C. and Duan, K.J. (1997) Immobilization of beta-fructofuranosidases from *Aspergillus* on methacrylamide-based polymeric beads for production of fructooligosaccharides. Biotechnol. Prog. *13*, 577–582.

Cobucci-Ponzano, B., Perugino, G., Trincone, A., Mazzone, M., Di Lauro, B., Giordano, A., Rossi, M. and Moracci, M. (2003) Applications in biocatalysis of glycosyl hydrolases from the hyperthermophilic archaeon *Sulfolobus solfataricus*. Biocatal. Biotransform. *21*, 215–221.

Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-Active Enzymes server at URL: http://afmb.cnrs-mrs.fr/CAZY/.

Daines, A.M., Maltman, B.A. and Flitsch, S.L. (2004) Synthesis and modifications of carbohydrates, using biotransformations. Curr. Opin. Chem. Biol. *8*, 106–113.

Davies, G.J., Charnock, S.J. and Henrissat, B. (2001) The enzymatic synthesis of glycosidic bonds: "Glycosynthases" and glycosyltransferases. Trends Glycosci. Glycotechnol. *13*, 105–120.

Devulapalle, K.S., Gómez de Segura, A., Ferrer, M., Alcalde, M., Mooser, G. and Plou, F.J. (2004) Effect of carbohydrate fatty acid esters on *Streptococcus sobrinus* and glucosyltransferase activity. Carbohydr. Res. *339*, 1029–1034.

Djouzi, Z., Andrieux, C., Pelenc, V., Somarriba, S., Popot, F., Paul, F., Monsan, P. and Szylit, O. (1995) Degradation and fermentation of alpha-gluco-oligosaccharides by bacterial strains from human colon: *in vitro* and *in vivo* studies in gnotobiotic rats. J. Appl. Bacteriol. *79*, 117–127.

Dols-Lafargue, M., Willemot, R.M., Monsan, P.F. and Remaud-Simeon, M. (2001) Factors affecting alpha,-1,2 glucooligosaccharide synthesis by *Leuconostoc mesenteroides* NRRL B-1299 dextransucrase. Biotechnol. Bioeng. *74*, 498–504.

Dols, M., Simeon, M.R., Willemot, R.M., Vignon, M.R. and Monsan, P.F. (1998) Structural characterization of the maltose acceptor-products synthesized by *Leuconostoc mesenteroides* NRRL B-1299 dextransucrase. Carbohydr. Res. *305*, 549–559.

Eggleston, G. and Cote, G.L. (2003) Oligosaccharides in food and agriculture. Washington: American Chemical Society.

Euzenat, O., Guibert, A. and Combes, D. (1997) Production of fructo-oligosaccharides by levansucrase from Bacillus subtilis C4. Process Biochem. *32*, 237–243.

Fernandez, R.C., Maresma, B.G., Juarez, A. and Martinez, J. (2004) Production of fructooligosaccharides by beta-fructofuranosidase from Aspergillus sp 27H. J. Chem. Technol. Biotechnol. *79*, 268–272.

Ghazi, I., Gómez de Segura, A., Fernández-Arrojo, L., Alcalde, M., Yates, M., Rojas-Cervantes, M.L., Plou, F.J. and Ballesteros, A. (2005) Immobilisation of fructosyltransferase from *Aspergillus aculeatus* on epoxy-activated Sepabeads EC for the synthesis of fructo-oligosaccarides. J. Mol. Catal. B: Enz. *35*, 19–27.

Gibson, G.R. and Ottaway, R.A. (2000) Prebiotics: New developments in functional foods. Oxford: Ed. Chandos Publishing.

Gonzalez-Munoz, F., Perez-Oseguera, A., Cassani, J., Jimenez-Estrada, M., Vazquez-Duhalt, R. and Lopez-Munguia, A. (1999) Enzymatic synthesis of fructosyl glycerol. J. Carbohydr. Chem. *18*, 275–283.

Goulas, A.K., Cooper, J.M., Grandison, A.S. and Rastall, R.A. (2004) Synthesis of isomalto-oligosaccharides and oligodextrans in a recycle membrane bioreactor by the combined use of dextransucrase and dextranase. Biotechnol. Bioeng. *88*, 778–787.

Gómez de Segura, A., Alcalde, M., Plou, F.J., Remaud-Simeon, M., Monsan, P. and Ballesteros, A. (2003) Encapsulation in LentiKats of dextransucrase from *Leuconostoc mesenteroides* NRRL B-1299, and its effect on product selectivity. Biocatal. Biotransform. *21*, 325–331.

Grizard, D. and Barthomeuf, C. (1999) Enzymatic synthesis and structure determination of NEO-FOS. Food Biotechnol. *13*, 93–105.

Hamilton, C.J. (2004) Enzymes in preparative mono- and oligo-saccharide synthesis. Natural Product Reports *21*, 365–385.

Hayashi, S., Yoshiyama, T., Fujii, N. and Shinohara, S. (2000) Production of a novel syrup containing neofructo-oligosaccharides by the cells of *Penicillium citrinum*. Biotechnol. Lett. *22*, 1465–1469.

Honda, Y. and Kitaoka, M. (2006) The first glycosynthase derived from an inverting glycoside hydrolase. J. Biol. Chem. *281*, 1426–1431.

Ichikawa, Y., Look, G.C., Wong, C.H. and Kajimoto, T. (1992) Synthesis of oligosaccharides using glycosyltransferases. J. Synth. Org. Chem. Japan *50*, 441–450.

Jeanes, A., Haynes, W.C., Wilham, C.A., Rankin, J.C., Melvin, E.H., Austin, M.J., Cluskey, J.E., Fisher, H.M., Tsuchiya, H.M. and Rist, C.E. (1954) Characterization and classification of dextrans from ninety-six strains of bacteria. J. Am. Chem. Soc. *76*, 5041–5052.

Kato, N., Suyama, S., Shirokane, M., Kato, M., Kobayashi, T. and Tsukagoshi, N. (2002) Novel alpha-glucosidase from Aspergillus nidulans with strong transglycosylation activity. Appl. Environ. Microbiol. *68*, 1250–1256.

Kilian, S., Kritzinger, S., Rycroft, C., Gibson, G. and du Preez, J. (2002) The effects of the novel bifidogenic trisaccharide, neokestose, on the human colonic microbiota. World J. Microbiol. Biotechnol. *18*, 637–644.

Kilian, S.G., Sutherland, F.C.W., Meyer, P.S. and duPreez, J.C. (1996) Transport-limited sucrose utilization and neokestose production by *Phaffia rhodozyma*. Biotechnol. Lett. *18*, 975–980.

Kim, M.G., Kim, C.H., Lee, J.S., Song, K.B. and Rhee, S.K. (2000) Synthesis of methyl beta-D-fructoside catalyzed by levansucrase from *Rahnella aquatilis*. Enz. Microb. Technol. *27*, 646–651.

Kim, T.K., Park, D.C. and Lee, Y.H. (1997) Synthesis of glucosyl-sugar alcohols using glycosyltransferases and structural identification of glucosyl-maltitol. J. Microbiol. Biotechnol. *7*, 310–317.

Kometani, T., Terada, Y., Nishimura, T., Takii, H. and Okada, S. (1994) Transglycosylation to hesperidin by cyclodextrin glucanotransferase from an alkalophilic *Bacillus* species in alkaline pH and properties of hesperidin glycosides. Biosci. Biotechnol. Biochem. *58*, 1990–1994.

Kren, V. and Thiem, J. (1997) Glycosylation employing bio-systems: from enzymes to whole cells. Chem. Soc. Rev. *26*, 463–473.

Kritzinger, S.M., Kilian, S.G., Potgieter, M.A. and du Preez, J. C. (2003) The effect of production parameters on the synthesis of the prebiotic trisaccharide, neokestose, by *Xanthophyllomyces dendrorhous* (*Phaffia rhodozyma*) Enz. Microb. Technol. *32*, 728–737.

Leemhuis, H. and Dijkhuizen, L. (2003) Engineering of hydrolysis reaction specificity in the transglycosylase cyclodextrin glycosyltransferase. Biocatal. Biotransform. *21*, 261–270.

Leemhuis, H., Rozeboom, H.J., Wilbrink, M., Euverink, G.J.W., Dijkstra, B. W. and Dijkhuizen, L. (2003) Conversion of cyclodextrin glycosyltransferase into a starch hydrolase by directed evolution: The role of alanine 230 in acceptor subsite+1. Biochemistry *42*, 7518–7526.

Mackenzie, L.F., Wang, Q.P., Warren, R.A.J. and Withers, S.G. (1998) Glycosynthases: Mutant glycosidases for oligosaccharide synthesis. J. Am. Chem. Soc. *120*, 5583–5584.

Macmillan, D. and Daines, A.M. (2003) Recent developments in the synthesis and discovery of oligosaccharides and glycoconjugates for the treatment of disease. Curr. Med. Chem. *10*, 2733–2773.

Mala, S., Dvorakova, H., Hrabal, R. and Kralova, B. (1999) Towards regioselective synthesis of oligosaccharides by use of alpha-glucosidases with different substrate specificity. Carbohydr. Res. *322*, 209–218.

Mala, S. and Kralova, B. (2000) Heterooligosaccharide synthesis catalyzed by alpha-glucosidase from *Bacillus stearothermophilus*. J. Mol. Catal. B: Enz. *10*, 617–621.

Malet, C. and Planas, A. (1998) From beta-glucanase to beta-glucansynthase: glycosyl transfer to alpha-glycosyl fluorides catalyzed by a mutant endoglucanase lacking its catalytic nucleophile. FEBS Lett. *440*, 208–212.

Martin, M.T., Alcalde, M., Plou, F.J., Dijkhuizen, L. and Ballesteros, A. (2001) Synthesis of maltooligosaccharides via the acceptor reaction catalyzed by cyclodextrin glycosyltransferases. Biocatal. Biotransform. *19*, 21–35.

Marx, S.P., Winkler, S. and Hartmeier, W. (2000) Metabolization of beta-(2,6)-linked fructose-oligosaccharides by different bifidobacteria. FEMS Microbiol. Lett. *182*, 163–169.

McNaught, A.D. (1997) International Union of Pure and Applied Chemistry and International Union of Biochemistry and Molecular Biology - Joint Commission on Biochemical Nomenclature - Nomenclature of carbohydrates (Recommendations 1996) (Reprinted from Pure Appl Chem, vol 68, pp. 1919–2008, 1996) San Diego: Academic Press Inc.

Meng, G.Y. and Fütterer, K. (2003) Structural framework of fructosyl transfer in *Bacillus subtilis* levansucrase. Nat. Struct. Biol. *10*, 935–941.

Monchois, V., Willemot, R.M. and Monsan, P. (1999) Glucansucrases: mechanism of action and structure-function relationships. FEMS Microbiol. Rev. *23*, 131–151.

Nguyen, Q.D., Rezessy-Szabo, J.M., Bhat, M.K. and Hoschke, A. (2005) Purification and some properties of beta-fructofuranosidase from *Aspergillus niger* IMI303386. Proc. Biochem. *40*, 2461–2466.

Olivares-Illana, V., Wacher-Rodarte, C., Le Borgne, S. and Lopez-Munguia, A. (2002) Characterization of a cell-associated inulosucrase from a novel source: A *Leuconostoc citreum* strain isolated from Pozol, a fermented corn beverage of Mayan origin. J. Ind. Microbiol. Biotechnol. *28*, 112–117.

Ozimek, L.K., van Hijum, S.A.F.T., van Koningsveld, G.A., van der Maarel, M.J.E.C., Geel-Schutten, G.H. and Dijkhuizen, L. (2004) Site-directed mutagenesis study of the three catalytic residues of the fructosyltransferases of *Lactobacillus reuteri* 121. FEBS Lett. *560*, 131–133.

Park, D.C., Kim, T.K. and Lee, Y.H. (1998) Characteristics of transglycosylation reaction of cyclodextrin glucanotransferase in the heterogeneous enzyme reaction system using extrusion starch as a glucosyl donor. Enz. Microb. Technol. *22*, 217–222.

Park, H.E., Park, N.H., Kim, M.J., Lee, T.H., Lee, H.G., Yang, J.Y. and Cha, J. (2003) Enzymatic synthesis of fructosyl oligosaccharides by levansucrase from *Microbacterium laevaniformans* ATCC 15953. Enz. Microb. Technol. *32*, 820–827.

Perugino, G., Cobucci-Ponzano, B., Rossi, M. and Moracci, M. (2005) Recent advances in the oligosaccharide synthesis promoted by catalytically engineered glycosidases. Adv. Synth. Catal. *347*, 941–950.

Planas, A. and Faijes, M. (2002) Glycosidases and glycosynthases in enzymatic synthesis of oligosaccharides. An overview. Afinidad *59*, 295–313.

Plou, F.J., Martin, M.T., Gomez de Segura, A., Alcalde, M. and Ballesteros, A. (2002) Glucosyltransferases acting on starch or sucrose for the synthesis of oligosaccharides. Can. J. Chem. *80*, 743–752.

Probert, H.M. and Gibson, G.R. (2002) Investigating the prebiotic and gas-generating effects of selected carbohydrates on the human colonic microflora. Lett. Appl. Microbiol. *35*, 473–480.

Robyt, J.F. (1996) Mechanism and action of glucansucrases. In Park, K.H., Robyt, J.F. and Choi, Y.D. (Eds.) Enzymes for Carbohydrate Engineering (pp. 1–22) Amsterdam: Elsevier Science B.V.

Robyt, J.F. and Eklund, S.H. (1983) Relative, quantitative effects of acceptors in the reaction of *Leuconostoc mesenteroides* B-512F dextransucrase. Carbohydr. Res. *121*, 279–286.

Robyt, J.F. and Walseth, T.F. (1978) Mechanism of dextransucrase action. 4. Mechanism of acceptor reactions of *Leuconostoc mesenteroides* B-512F dextransucrase. Carbohydr. Res. *61*, 433–445.

Roujeinikova, A., Raasch, C., Burke, J., Baker, P.J., Liebl, W. and Rice, D.W. (2001) The crystal structure of *Thermotoga maritima* maltosyltransferase and its implications for the molecular basis of the novel transfer specificity. J. Mol. Biol. *312*, 119–131.

Sangeetha, P.T., Ramesh, M.N. and Prapulla, S.G. (2005) Fructooligosaccharide production using fructosyl transferase obtained from recycling culture of *Aspergillus oryzae* CFR 202. Process Biochem. *40*, 1085–1088.

Sanz-Aparicio, J., Hermoso, J.A., Martinez-Ripoll, M., Lequerica, J.L. and Polaina, J. (1998) Crystal structure of beta-glucosidase A from *Bacillus polymyxa*: Insights into the catalytic activity in family 1 glycosyl hydrolases. J. Mol. Biol. *275*, 491–502.

Scigelova, M., Singh, S. and Crout, D.H.G. (1999) Glycosidases- a great synthetic tool. J. Mol. Catal. B: Enz. *6*, 483–494.

Shin, H.T., Baig, S.Y., Lee, S.W., Suh, D.S., Kwon, S.T., Lim, Y.B. and Lee, J.H. (2004) Production of fructo-oligosaccharides from molasses by *Aureobasidium pullulans* cells. Biores. Technol. *93*, 59–62.

Shoda, S., Izumi, R. and Fujita, M. (2003) Green process in glycotechnology. Bull. Chem. Soc. Japan *76*, 1–13.

Simmering, R. and Blaut, M. (2001) Pro- and prebiotics - the tasty guardian angels? Appl. Microbiol. Biotechnol. *55*, 19–28.

Steinbchel, R.A. and Rhee, S.K.I. (2005) Polysaccharides and polyamide in the food industry: properties, production and patents. New York: Wiley-VCH.

Straathof, A.J.J., Kieboom, A.P.G. and Van Bekkum, H. (1986) Invertase-catalyzed fructosyl transfer in concentrated solutions of sucrose. Carbohydr. Res. *146*, 154–159.

Tambara, Y., Hormaza, J.V., Perez, C., Leon, A., Arrieta, J. and Hernandez, L. (1999) Structural analysis and optimised production of fructo-oligosaccharides by levansucrase from *Acetobacter diazotrophicus* SRT4. Biotechnol. Lett. *21*, 117–121.

Tanriseven, A. and Aslan, Y. (2005) Immobilization of Pectinex Ultra SP-L to produce fructooligosaccharides. Enz. Microb. Technol. *36*, 550–554.

Thiem, J. (1995) Applications of enzymes in synthetic carbohydrate-chemistry. FEMS Microbiol. Rev. *16*, 193–211.

Tonkova, A. (1998) Bacterial cyclodextrin glucanotransferase. Enz. Microb. Technol. *22*, 678–686.

Trujillo, L.E., Arrieta, J.G., Dafhnis, F., Garcia, J., Valdes, J., Tambara, Y., Perez, M. and Hernandez, L. (2001) Fructo-oligosaccharides production by the *Gluconacetobacter diazotrophicus* levansucrase expressed in the methylotrophic yeast *Pichia pastoris*. Enz. Microb. Technol. *28*, 139–144.

Tungland, B.C. (2003) Fructooligosaccharides and other fructans: Structures and occurrence, production, regulatory aspects, food applications, and nutritional health significance. In Eggleston, G. and Cote, G.L. (Eds.) Oligosaccharides in Food and Agriculture (pp. 135–152). Washington: American Chemical Society.

Tuohy, K.M., Rouzaud, G.C.M., Bruck, W.M. and Gibson, G.R. (2005) Modulation of the human gut microflora towards improved health using prebiotics - Assessment of efficacy. Curr. Pharm. Des. *11*, 75–90.

van der Veen, B.A., van Alebeek, G.J.W.M., Uitdehaag, J.C.M., Dijkstra, B.W. and Dijkhuizen, L. (2000) The three transglycosylation reactions catalyzed by cyclodextrin glycosyltransferase from *Bacillus circulans* (strain 251) proceed via different kinetic mechanisms. Eur. J. Biochem. *267*, 658–665.

Velasco, J. and Adrio, J.L. (2002) Microbial enzymes for food-grade oligosaccharide biosynthesis. In Barredo, J.L. (Ed.) Micro-organisms for health care, food and enzyme production, pp. 173–186, Kerala, India: Research Signpost.

Vic, G. and Crout, D.H.G. (1995) Synthesis of allyl and benzyl beta-D-glucopyranosides, and allyl beta-D-galactopyranoside from D-glucose or D-galactose and the corresponding alcohol using almond beta-D-glucosidase. Carbohydr. Res. *279*, 315–319.

Vic, G., Hastings, J.J. and Crout, D.H.G. (1996) Glycosidase-catalysed synthesis of glycosides by an improved procedure for reverse hydrolysis: Application to the chemoenzymatic synthesis of galactopyranosyl-(1→4)-O-alpha-galactopyranoside derivatives. Tetrahedron-Asymm. *7*, 1973–1984.

SECTION B

PEPTIDASES

CHAPTER 10

AN INTRODUCTION TO PEPTIDASES AND THE *MEROPS* DATABASE

NEIL D. RAWLINGS*, FRASER R. MORTON AND ALAN J. BARRETT
Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, United Kingdom
**ndr@sanger.ac.uk*

1. INTRODUCTION

Peptidases are enzymes that hydrolyse peptide bonds. They are necessary for the survival of all living creatures, and they are encoded by about 2% of genes in all kinds of organisms. It has been estimated that 14% of the five hundred human peptidases are under investigation as drug targets (Southan, 2001). Peptidases are important for many biological processes including digestion of food proteins, recycling of intracellular proteins, the blood coagulation cascade, antigen presentation, and activation of a variety of proteins, including enzymes, peptide hormones, and neurotransmitters. We calculate that 18% of all proteins in the SwissProt protein sequence database are annotated as undergoing post-translational proteolytic processing during maturation.

There are many industrial uses for peptidases, though often mixtures rather than purified enzymes are used. The earliest were in cheese making, initially using plant juices to clot milk (according to Homer in *The Iliad,* fig juice, which contains ficain, C01.006, was used) and then 'rennet', animal stomach contents that contain chymosin (A01.006). Peptidases are also used to tenderize meat, clarify beers and enhance the flavours of cheeses and pet foods. Peptidases are used in the leather industry to remove hair, and make the leather more supple ("bating" and "soaking" the leather); however many of these are proprietary products for which the sequences and organisms of origin are not public. Peptidases are also widely used in cleaning materials, such as biological washing powders and contact lens cleaning fluid. Besides being the targets of drugs, peptidases are used in medicine to remove gastrointestinal parasites (anthelminthics), removal of dead skin from burn patients (debridement), determination of blood groups, and for relief

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 161–179

of back pain by digesting the cartilage content of herniated intervertebral discs (chemonucleolysis). Peptidases are also widely used as reagents in the laboratory for limited proteolysis of proteins, and for generating the peptides required for protein sequencing. A more recent laboratory use for peptidases of restricted specificity has been in the processing of recombinant fusion proteins. For example, the NIa endopeptidases from tobacco etch virus (C04.004) processes the viral polyprotein at a specific site: ENLYFQ+G/S. Introducing a fragment of nucleotide sequence that encodes such a site in the vector between that coding for the proteins (or between that coding for polyHis and a protein) enables individual proteins to be separated from the fusion product by addition of the endopeptidase (Kapust *et al*., 2000). Examples of uses of peptidases are shown in Table 1; the review by Rao *et al*., 1998 provides many more examples (Rao *et al*., 1998).

Peptidases are thus an exceptionally important group of enzymes in biology, medical research and biotechnology. Since the regulation of the activities of peptidases is obviously crucial, the hundreds of proteins that inhibit them are equally relevant.

In this chapter we will describe some of the terms relevant to peptidases, then discuss the various classification methods, with particular reference to that employed by the *MEROPS* database. In the *MEROPS* database a unique identifier is given to each different peptidase, and whenever a peptidase is mentioned in the text, this identifier will also be given. This will enable the reader to obtain further information about each peptidase from the *MEROPS* database.

2. PEPTIDASE

Peptidase is the term recommended in the Enzyme Nomenclature of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB) and by the Human Gene Nomenclature Committee, as well as the *MEROPS* database for any protein that causes the hydrolysis of peptide bonds. The fact that "peptidase" already forms the root of the names of the many different sub-types of peptidases: aminopeptidase, carboxypeptidase, and so on (see below), leads to a very rational and intuitive system of terminology. It is applicable to the endopeptidases that act on the internal bonds in proteins and large polypeptides as well as to the oligopeptidases and exopeptidases that act primarily on smaller substrates. Peptidases are also known colloquially as proteases, proteinases and proteolytic enzymes.

2.1. Catalytic Type

The *catalytic type* of a peptidase relates to the chemical groups responsible for its catalysis of peptide bond hydrolysis. The six specific catalytic types that are recognised are the serine, threonine, cysteine, aspartic, glutamic and metallo- peptidases. In peptidases of serine, threonine and cysteine type, the catalytic nucleophile is the reactive group of an amino acid side chain, either a hydroxyl group (serine

Table 1. Commercial uses of peptidases Peptidases are listed in order of *MEROPS* identifier (MERID)

MERID	Name	Commercial uses
A01.001	pepsin A	Protein sequencing.
A01.006	chymosin	Cheese making.
A01.013	mucorpepsin	Cheese making.
A01.017	endothiapepsin	Cheese making.
A01.020	phytepsin	Contributes to the malting of cereal grains. Forms from *Centauria* sp. and *Cynara* sp. have been used to coagulate milk in cheese making.
A26.001	omptin	An engineered variant of omptin has been proposed as a specific endopeptidase for the cleavage of recombinant fusion proteins.
C01.001	papain	Used in food processing.
C01.002	chymopapain	Contact-lens cleaning fluid. Stem cell isolation. Treatment of herniated intervertebral disk.
C01.004	glycyl endopeptidase	Protein sequencing.
C01.005	stem bromelain	Blood group determination.
C01.006	ficain	Blood group determination.
C01.026	ananain	Burn debridement.
C01.086	aminopeptidase C	Cheese making.
C04.001	nuclear-inclusion-A endopeptidase (plum pox virus)	Processing of recombinant fusion proteins.
C04.003	tobacco vein mottling virus-type NIa endopeptidase	Processing of recombinant fusion proteins.
C04.004	tobacco etch virus NIa endopeptidase	Processing of recombinant fusion proteins.
C15.001	pyroglutamyl-peptidase I (prokaryote)	Removal of pyroglutamyl groups from peptides, in protein sequencing. Laboratory use in identification of group A streptococci and enterococci.
M09.001	microbial collagenase (*Vibrio* sp.)	Several proposed uses, including tissue cell dispersion, burn debridement.
M09.002	collagenase colA	Many uses, including tissue cell dispersion.
M11.001	gametolysin	Use in preparation of *Chlamydomonas* protoplasts.
M12.066	flavastacin	Specificity is similar to that of an endopeptidase sold for protein sequencing as 'endoproteinase Asp-N'.
M12.151	ecarin	Laboratory use in haematology.
M12.158	russellysin	Laboratory use in haematology.
M20.001	glutamate carboxypeptidase	Glutamate carboxypeptidase has been the subject of considerable research in experimental prodrug strategies for cancer therapy including 'antibody-directed enzyme prodrug therapy' (ADEPT).
M22.001	O-sialoglycoprotein endopeptidase	Reagent in study of mammalian cell surface sialoglycoproteins.

(Continued)

Table 1. (Continued)

MERID	Name	Commercial uses
M23.004	lysostaphin	An agent for lysis of staphylococcal cell walls in the laboratory.
M27.002	bontoxilysin	As Botox, has therapeutic use for local paralysis of neuromuscular function, as in strabismus.
M35.002	deuterolysin	Taste-forming factor in soy sauce.
M42.001	glutamyl aminopeptidase (bacterium)	Contributes to maturation of cheese.
M72.001	peptidyl-Asp metalloendopeptidase	Reagent in protein sequencing.
M9A.008	tryptophanyl aminopeptidase	Use in L-tryptophan manufacture.
M9G.055	Dispase	Used in tissue cell dispersion.
S01.001	chymotrypsin A (cattle)	Protein sequencing. Removal of allergens from milk protein hydrolysates.
S01.151	Trypsin 1	Protein sequencing. Preparation of bacterial media. Bating leather.
S01.156	enteropeptidase	Reagent for cleavage of recombinant fusion proteins.
S01.176	batroxobin	Used as benign defibrinating agent
S01.177	crotalase	Used as benign defibrinating agent.
S01.178	Ancrod	Used as benign defibrinating agent.
S01.216	coagulation factor Xa	Reagent for cleavage of recombinant fusion proteins.
S01.217	thrombin	Reagent for cleavage of recombinant fusion proteins.
S01.261	streptogrisin A	A component of Pronase.
S01.262	streptogrisin B	A component of Pronase.
S01.265	streptogrisin C	A component of Pronase.
S01.266	streptogrisin D	A component of Pronase.
S01.267	streptogrisin E	A component of Pronase.
S01.269	glutamyl endopeptidase I	Used in selective hydrolysis of proteins.
S01.280	lysyl endopeptidase (bacteria)	Used in amino acid sequencing of proteins.
S08.001	subtilisin Carlsberg	Forms of subtilisin are widely used commercially. Alcalase (from *Bacillus licheniformis*), Esperase (from *Bacillus*) and Maxatase (from *Bacillus*) are commercial names for peptidases used in biological washing powders. Alcalase is also used in the food industry to process whey and in the production of pet food.
S08.019	lactocepin I	Role in digestion of caseins by lactobacilli in cheese making.
S08.056	cuticle-degrading endopeptidase	Cuticle-degrading endopeptidase contributes to the effectiveness of organisms used in the biocontrol of insect and nematode pests.
S08.071	furin	Proposed for use in processing of recombinant proteins.
S10.016	Carboxypeptidase S1	Used to enhance flavours in foods; commercially available as Flavourzyme.

and threonine peptidases) or a sulfhydryl group (cysteine peptidases). In aspartic and metallo- peptidases, the nucleophile is commonly an activated water molecule. In aspartic peptidases, the water molecule is directly bound by the side chains of aspartic residues. In metallopeptidases, one or two divalent metal ions hold the water molecule in place, and charged amino acid side chains are ligands for the metal ions. The metal is most commonly zinc, but may also be cobalt, manganese or copper. A single metal ion is usually bound by three amino acid ligands. The activated water molecule is a fourth metal ligand, and the metal is described as "tetrahedrally co-ordinated". Where two metal ions are present, each is tetrahedrally co-ordinated, so that two activated water molecules are bound, and one amino acid residue ligates both metals. The glutamic peptidases (all in the small family G1) were recognised only in 2005 (Kataoka *et al.*, 2005), and much remains to be learned about their catalytic mechanisms, but they seem to employ a Glu/Gln catalytic dyad. Just a few peptidases are still of unknown catalytic type.

2.2. Active Site

Crystallographic structures of peptidases show that the active site is commonly located in a groove on the surface of the molecule between adjacent structural domains, and the substrate specificity is dictated by the properties of binding sites arranged along the groove on one or both sides of the catalytic site that is responsible for hydrolysis of the bond cleaved (the **scissile bond**). Besides the nucleophile, other residues are important for catalysis and maintaining the structure of the active site. The active site residues are very well conserved between all the active peptidases within a family.

In general terms, cleavage of a peptide bond has been described as an example of an acid/base reaction, in which the charged nucleophile is the proton donor and a residue known as the general base is the proton acceptor. In serine and cysteine peptidases the general base is often a histidine, but can be a lysine (*e.g.* signal peptidase I, S26.001 and endopeptidase La, S16.001). When the general base is a histidine, usually a third residue orientates the imidazolium ring of the histidine and helps charge one of the nitrogen atoms in the ring. In many serine peptidases this third member of the **catalytic triad** is an aspartate, for example in chymotrypsin (S01.001), subtilisin (S08.001) and carboxypeptidase Y (S10.001). In assemblin (S21.001) the third residue is a second histidine, and in D-Ala-D-Ala carboxypeptidase A (S11.001) it is a second serine. Exceptionally, the serine peptidases omptin (S18.001) and eukaryote signal peptidase (S26.010) have a Ser/His catalytic dyad only. In cysteine peptidases the third member of the triad may be asparagine (*e.g.* papain, C01.001), aspartate (*e.g.* deubiquitinating peptidase Yuh1, C12.001) or glutamate (*e.g.* adenovirus endopeptidase, C05.001). There are many cysteine peptidases which have only a Cys/His dyad, however.

In serine and cysteine peptidases, a fourth residue is often important because it helps stabilize the transitional acyl-intermediate that forms between the peptidase and the substrate as a first stage of catalysis. A residue forms a hydrogen bond

with the negatively charged oxygen atom, and this catalytic subsite is known as the **oxyanion hole**. In chymotrypsin this fourth important residue is glycine, in subtilisin it is asparagine and in papain it is glutamine.

Some peptidases appear to have only one catalytic residue, which is the N-terminal residue. These are known as **N-terminal nucleophile (Ntn) hydrolases**. All known threonine peptidases are Ntn-hydrolases, but there are also some serine peptidases (*e.g.* penicillin G acylase precursor, S45.001) and cysteine peptidases (*e.g.* penicillin V acylase precursor, C59.001), that are autolytic peptidases. In Ntn-hydrolases, the N-terminal amino group is thought to function as the general base.

Full descriptions of the catalytic mechanisms of serine, cysteine and threonine peptidases have been provided by Polgar (2004) (Polgar, 2004a; Polgar, 2004b).

No residues other than the aspartates are known to be involved in catalysis by the aspartic peptidases (James, 2004). In metallopeptidases other residues have been shown by mutation studies to be essential, but exactly what their roles may be is controversial (Auld, 2004). A glutamate is important for activity in all the metallopeptidases that carry the HEXXH zinc-binding motif (*e.g.* thermolysin, M04.001), as well as carboxypeptidase A (M14.001). In metallopeptidases that have two catalytic metal ions, two residues are essential, often a glutamate and an aspartate (*e.g.* glutamate carboxypeptidase, M20.001).

2.3. Terminology of Peptidase Specificity: Schechter and Berger Nomenclature

The specificity of a peptidase is described by use of a conceptual model in which each specificity subsite is able to accommodate the side chain of a single amino acid residue. The sites are numbered from the catalytic site, S1, S2 ... Sn towards the N-terminus of the substrate, and S1′, S2′...Sn′ towards the C-terminus. The residues they accommodate are numbered P1, P2 ... Pn, and P1′, P2′... Pn′, respectively, as follows:

Substrate: - P3 - P2 - P1+ P1′ - P2′ - P3′ -

Enzyme: - S3 - S2 - S1 * S1′ - S2′ - S3′ -

In this representation the catalytic site of the enzyme is marked * and the scissile bond is indicated by the symbol +. This system is based on one that was first used by Schechter and Berger in relation to papain (Schechter and Berger, 1967).

3. CLASSIFICATION OF PEPTIDASES

A landmark in the development of any field of study is the appearance of a sound system of nomenclature and classification for the objects with which it deals. The introduction of the Linnaean system for naming and classifying organisms in the

eighteenth century and the invention of a system of nomenclature for enzymes in the 1950's were such key events, and their value has been obvious. Both nomenclature and classification are vitally important for information-handling, allowing people to communicate efficiently, knowing that they are talking about the same thing, and to store and retrieve information unambiguously. A good system also serves to highlight important questions and thus prompts new discoveries. Three useful methods of grouping peptidases are currently in use: i) by the chemical mechanism of catalysis; ii) by the details of the reaction catalysed; iii) by molecular structure and homology. Each of these is described below in more detail.

3.1. Peptidases Grouped by the Chemical Mechanism of Catalysis

In 1960 the seminal paper of Hartley (Hartley, 1960) initiated a sequence of developments that has now provided the peptidase community with the very useful concept of catalytic type. The system of catalytic types (as described above) has great strengths, but it also has limitations that need to be recognised. It is a strength that every serine peptidase contains a serine residue that acts as the nucleophile at the heart of the catalytic site, and as a result many are affected by generic inhibitors of serine peptidases. But the serine peptidases include many very different molecular structures and catalytic mechanisms. Moreover, they are by no means all homologues of each other, so an expression like "the serine peptidase family" has little meaning.

3.2. Peptidases Grouped by the Kinds of Reaction they Catalyse

In a sense, all peptidases catalyse the same reaction: hydrolysis of a peptide bond. But they are selective for the position of the peptide bond in the substrate, for the amino acid residues near the scissile bond, and for other characteristics of the substrate that are still not understood. The terms used to describe different specificities are explained below and shown diagrammatically in Fig. 1.

3.2.1. Endopeptidases

An endopeptidase hydrolyses internal, alpha-peptide bonds in a polypeptide chain, tending to act away from the N-terminus or C-terminus. Examples of endopeptidases are chymotrypsin (S01.001; (Graf *et al.*, 2004)), pepsin (A01.001; (Tang, 2004)) and papain (C01.001; (Menard *et al.*, 2004a)). Some endopeptidases act only on substrates smaller than proteins, and these are termed **oligopeptidases**. An example of an oligopeptidase is thimet oligopeptidase (M03.001; (Barrett *et al.*, 2004)). Endopeptidases initiate the digestion of food proteins, generating new N- and C-termini that are substrates for the exopeptidases that complete the process. Endopeptidases also process proteins by limited proteolysis. Examples are the removal of signal peptides from secreted proteins (*e.g.* signal peptidase I, S26.001; (Dalbey, 2004)) and the maturation of precursor proteins (*e.g.* enteropeptidase, S01.156, (Sadler, 2004); furin, S08.071, (Creemers *et al.*, 2004)). A very few

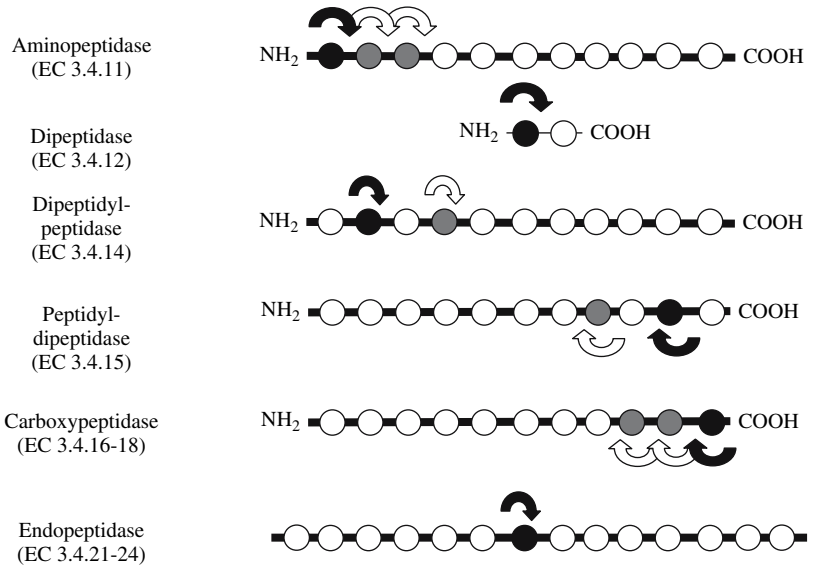


Figure 1. **Classification of peptidases by reaction catalysed.** Peptides are represented as beads on a string, with each bead representing an amino acid and the string representing the peptide bonds. N- ("NH_2") and C- ("COOH") termini are indicated. Black arrows show the first cleavage and white arrows show subsequent cleavages. For the first cleavage, the amino acid(s) to which specificity is mainly directed is shown in black and for subsequent cleavages in grey

endopeptidases act at a fixed distance from one terminus of the substrate, an example being mitochondrial intermediate peptidase (M03.006; (Isaya, 2004)), which releases an N-terminal octapeptide. This octapeptide is the second of two N-terminal targeting signals of nuclear-encoded proteins that are imported into the mitochondrion. In the nomenclature of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB) endopeptidases are allocated to sub-subclasses EC 3.4.21, EC 3.4.22, EC 3.4.23, EC 3.4.24 and EC 3.4.25 for serine-, cysteine-, aspartic-, metallo- and threonine-type endopeptidases, respectively (NC-IUBMB, 1992).

3.2.2. Omega-peptidases

The omega-peptidases form the second group of peptidases that have no requirement for a free N-terminus or C-terminus in the substrate. Despite their lack of requirement for a charged terminal group, they often act close to one terminus or the other, and are thus totally distinct from endopeptidases. Some hydrolyse peptide bonds that are not alpha-bonds; that is, they are isopeptide bonds, in which one or both of the amino and carboxyl groups are not directly attached to the alpha-carbon of the parent amino acid. The omega-peptidases are a varied assortment of enzymes, including ubiquitinyl hydrolases (*e.g.* ubiquitinyl hydrolase-L3, C12.003; (Wilkinson, 2004)), pyroglutamyl peptidases (C15.010, (Dando, 2004); M01.008,

(Bauer, 2004)) and gamma-glutamyl hydrolase (C26.001; (Chave *et al*., 2004)). The omega-peptidases are placed in sub-subclass EC 3.4.19 by NC-IUBMB.

3.2.3. Exopeptidases

The exopeptidases require a free N-terminal amino group, C-terminal carboxyl group or both, and hydrolyse a bond not more than three residues from the terminus. The exopeptidases are further divided into aminopeptidases, carboxypeptidases, dipeptidyl-peptidases, peptidyl-dipeptidases, tripeptidyl-peptidases and dipeptidases. There are no known exopeptidases that are aspartic or glumatic peptidases.

Aminopeptidases. An aminopeptidase liberates a single amino acid residue from the unblocked N-terminus of its substrate: Xaa + peptide (or Xaa + $(Xaa)_n$). Examples are aminopeptidase N (M01.001; (Turner, 2004)) and aminopeptidase C (C01.086; (Chapot-Chartier, 2004)). Aminopeptidases form sub-subclass EC 3.4.11 in the NC-IUBMB scheme.

Dipeptidases. A dipeptidase hydrolyses a dipeptide, and requires that both termini be free: Xaa + Xaa. Examples are dipeptidase A (C69.001; (Dudley and Steele 2004,) and membrane dipeptidase (M19.001; (Hooper, 2004a)). Dipeptidases form sub-subclass EC 3.4.13 in the NC-IUBMB scheme.

Dipeptidyl-peptidases. A dipeptidyl-peptidase is so-called because it hydrolyses a dipeptidyl bond, *i.e.* it releases an N-terminal dipeptide from its substrate: dipeptide + peptide (*i.e.* $(Xaa)_2 + (Xaa)_n$), and that being the case, the term dipeptidyl-peptidase (short for 'dipeptidyl-peptide hydrolase') is clearly appropriate. These enzymes are sometimes erroneously called aminopeptidases or dipeptidases. Examples are dipeptidyl-peptidase I (C01.070; (Turk *et al*., 2004)) and dipeptidyl-peptidase III (M49.001; (Chen *et al*., 2004)). Dipeptidyl-peptidases, together with tripeptidyl-peptidases, form sub-subclass EC 3.4.14 in the NC-IUBMB scheme.

Tripeptidyl-peptidases. A tripeptidyl-peptidase hydrolyses a tripeptidyl bond, releasing a tripeptide from the N-terminus of its substrate: tripeptide + peptide (*i.e.* $(Xaa)_3 + (Xaa)_n$), and again, this explains the name. Examples are tripeptidyl-peptidase I (S53.003; (Sohar *et al*., 2004)) and tripeptidyl-peptidase II (S08.090; (Tomkinson, 2004)). Tripeptidyl peptidases, together with dipeptidyl-peptidases, form sub-subclass EC 3.4.14 in the NC-IUBMB scheme.

Peptidyl-dipeptidases. A peptidyl-dipeptidase hydrolyses a dipeptide from the C-terminus of its substrate: peptide + dipeptide (*i.e.* $(Xaa)_n + (Xaa)_2$), and this explains the name. An example is peptidyl-dipeptidase A (XM02-001; (Corvol *et al*., 2004)). Peptidyl-dipeptidases form sub-subclass EC 3.4.15 in the NC-IUBMB scheme.

Carboxypeptidases. A carboxypeptidase hydrolyses a single residue from the unblocked C-terminus of its substrate: peptide + Xaa (or more precisely: $(Xaa)_n$ + Xaa). Examples are carboxypeptidase A1 (M14.001; (Auld, 2004)),

cathepsin X (C01.013; (Menard *et al*., 2004b)) and carboxypeptidase Y (S10.001, (Mortensen *et al*., 2004)). Carboxypeptidases form sub-subclasses EC 3.4.16-18 in the NC-IUBMB scheme, being divided by catalytic type.

Other terms. Several other terms have been introduced for peptidases. The commonest of these extra terms is **tripeptidase**. A tripeptidase is a peptidase that is known only to degrade a tripeptide; however, the known tripeptidases are specialized aminopeptidases that release an N-terminal amino acid and a dipeptide and are consequently also known as "aminotripeptidases". An example is peptidase T (M20.003; (Miller *et al*., 2004)).

3.2.4. Limitations of classification by reaction

There are several limitations to this classification. By far the most important is that the classification does not reflect evolutionary relationships between the peptidases, because related peptidase can have very different substrate specificities and unrelated peptidases can have virtually identical substrate specificities, and thus be included in the same entry in the NC-IUBMB scheme. Endopeptidases are difficult to classify by this system because it is difficult to describe the reaction catalysed. For both carboxypeptidases and endopeptidases, catalytic type has been used to subdivide entries, even though substrate preference has little to do with catalytic type. This is inconsistent with the other sub-subclasses which also contain peptidases of different catalytic types.

3.3. Peptidases Grouped by Molecular Structure and Homology

The classification of peptidases by molecular structure and homology is the newest of the three methods, because it depends on the availability of data for amino acid sequences and three-dimensional structures in quantities that were realised only in the early 1990s. In 1993, Rawlings and Barrett described a system in which individual peptidases were assigned to families, and the families were grouped in clans (Rawlings *et al*., 1993). This scheme was developed to provide the structure of the *MEROPS* database, and has been extended to include the proteins that inhibit peptidases (Rawlings *et al*., 2004). The URL of the *MEROPS* database is: http://merops.sanger.ac.uk. The description below relates specifically to the way the classification of individual peptidases and inhibitors by molecular structure and homology is implemented in the *MEROPS* database.

3.3.1. Individual peptidases

Any one peptidase is expected to occur in many species of organisms, and these are known as **species variants**. Criteria we use to recognize the species variants of a single peptidase are as follows:

i) They have similar properties as enzymes, showing the same types and specificities of catalytic activity, pH optima and sensitivity to inhibitors. Where

biochemical data are unavailable, there are no differences in the protein sequences that would be predicted to result in differences in specificity.

ii) They have similar amino acid sequences throughout the length of the polypeptide encoded by the open reading frame.

iii) An evolutionary tree for the peptidase units shows that the protein sequences have diverged at the same time as the organisms in which they occur. An earlier divergence would imply that they are separate enzymes and not orthologues.

A single peptidase may include products of the allelic variants of a single gene and variants resulting from post-translational modification, and it may be expressed in different tissues or different stages of an organism's development. For each peptidase a single representative form termed the *holotype* is recognised. It is analogous to the *type peptidase* or *type inhibitor* at the family and clan levels of the classification.

Each individual peptidase is given a *MEROPS* identifier that is formed by concatenation of the three-character identifier of the family to which the peptidase belongs, a point, and a three-figure number. For example, the identifier of chymotrypsin, the type peptidase in family S1, is S01.001. A peptidase is considered to merit the assignment of an identifier when knowledge of it includes one or more amino acid sequences and information about substrate specificity or biological function. A satisfactory name is also very helpful. Over 2000 individual peptidases and over 500 inhibitors were recognised in Release 7.2 of the *MEROPS* database.

There are some peptidases that we have to treat as **unsequenced peptidases** because the available amino acid sequence data (if any) are insufficient to allow us to assign the peptidase to a family. In order to be able to present data for these peptidases we have created a series of special *MEROPS* identifiers in which the family name part of the identifier is replaced by a code that indicates only the catalytic type and the kind of peptidase activity. The first character of this shows the catalytic type as in a family identifier, the second character is always 9, and the third is a letter that indicates the kind of peptidase activity: 'A' for aminopeptidase, 'B' for dipeptidase, 'C' for dipeptidyl-peptidase, 'D' for peptidyl-dipeptidase, 'E' for carboxypeptidase, 'F' for omega peptidase and 'G' for endopeptidase. An example would be the *MEROPS* ID M9A.007 for Xaa-Trp aminopeptidase (Hooper, 2004b). As soon as fuller sequence data appear for an unsequenced peptidase we assign it a normal *MEROPS* ID.

3.3.2. Unassigned peptidases

In the past a protein was characterized first and the amino acid sequence came later, but with the advance of methods in sequence determination, especially the ability to sequence whole genomes, the reverse is now true and determination of a sequence commonly precedes characterization of the protein. It can be very difficult to discover the physiological substrates of a peptidase, because some peptidases have such restricted specificity that only a single protein substrate is cleaved (*e.g.* renin, A01.007, which only cleaves angiotensinogen (Suzuki *et al.*, 2004)). There are now many peptidase homologues that cannot be assigned to any *MEROPS* identifier

because the sequence is too different from that of any holotype. Consequently, we describe such a protein as an **unassigned homologue**, and a *MEROPS* identifier will be created when the biochemical characterization comes along.

3.3.3. Non-peptidase homologues

For many peptidase families we now know of homologues that are not peptidases, for example the S1 family includes azurocidin, haptoglobins and protein Z. In all of these cases at least one residue of the catalytic triad has been replaced. There are several homologues in family M12 wherein the zinc ligands have been replaced, and these are unable to bind zinc and are not peptidases. Such a protein is termed a **non-peptidase homologue**.

In order to classify every human and mouse non-peptidase homologues we have used some special *MEROPS* identifiers for these species. These all have a nine as the first digit after the dot. Examples are haptoglobin-1 (S01.972), mitochondrial processing peptidase alpha subunit (M16.971) and proteasome alpha 1 subunit (T01.976).

There are also some peptidase homologues that possess all the active site residues and/or metal ligands which are not known to cleave peptide bonds but are known to catalyse other reactions. An example is acetylornithine deacetylase which is a non-peptidase homologue in family M20. Another member of M20 from bacteria, succinyl-diaminopimelate desuccinylase (M20.010), was thought to be a non-peptidase homologue possessing all components of the active site, including the metal ligands, but has now been shown to act as a peptidase when the zinc is replaced by manganese (Broder *et al*., 2003).

Some non-peptidase homologues are enzymes of other kinds. An example is dienelactone hydrolase (EC 3.1.1.45), a member of family S9 that has the catalytic serine replaced by cysteine.

3.3.4. Peptidase unit

The *peptidase unit* is that part of the protein sequence that is directly responsible for peptidase activity, as far as it is known to *MEROPS*. In the simplest case, this is that part of the sequence that aligns with the smallest mature peptidase molecule in the family. In structural terms, the peptidase unit consists of two subdomains with the active site in the cleft between the domains.

Many peptidases and their precursors are **chimeric proteins** containing non-peptidase domains at the N- or C-terminus, or even inserted into the middle of the peptidase unit (in such a circumstance, the peptidase unit is described as **interrupted**, and each inserted domain is known as **nested**). For example, procollagen C-peptidase (M12.005) is a chimeric protein that contains a catalytic domain related to that of astacin, but also contains segments that are clearly homologous to non-catalytic parts of the complement components C1r and C1s, which are in the chymotrypsin family (Rawlings *et al*., 1990). The procollagen endopeptidase is placed in the family of astacin (M12), and not in that of chymotrypsin (S1). All

members of subfamily S41B have interrupted peptidase units, containing a nested PDZ domain (Ponting *et al*., 1999).

In some families even the smallest mature peptidase can be seen to be a multidomain protein by the presence of a segment that is homologous to a known non-peptidase domain found in other proteins. An example is family S16, in which all peptidases have an N-terminal ATPase domain (Vasilyeva *et al*., 2002). Such a domain is excluded from the peptidase unit. Since it is the case that for most peptidases the limits of the peptidase unit are inferred indirectly from a multiple sequence alignment, they can be refined from time to time as new data become available. Examples of peptidase units are shown in Fig. 2.

3.3.5. Compound and complex peptidases

The *MEROPS* classification of peptidases is a classification of peptidase units, and the great majority of proteins with peptidase activity contain only a single peptidase unit. But occasionally it happens that a single protein molecule contains several peptidase units. Such a molecule clearly requires special treatment because no single location in the classification is right for it. We term such a peptidase a **compound peptidase**. There are also multi-subunit peptidase molecules that

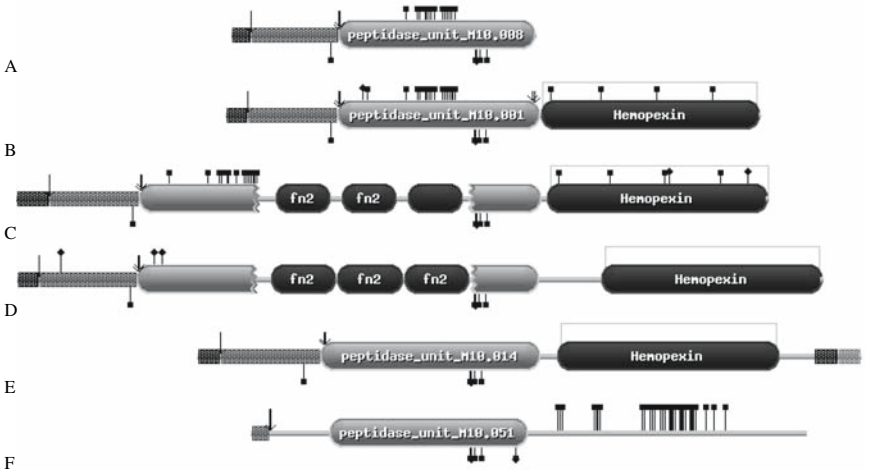


Figure 2. **Examples of peptidase units from family M10.** The images are proportional to the sequence length. Domains are shown as rounded rectangles; peptidase units are shown in grey and other domains in black. Small rectangles show signal peptides and transmembrane domains (black), activation peptides (dark grey) and cytoplasmic regions (light grey). Features shown on the top edge are cleavage positions (arrows), structural metal ligands (black squares), carbohydrate attachment sites (black diamonds) and disulfide bridges (grey lines). Features shown on the bottom edge are catalytic metal ligands (black squares) and active site residues (black diamonds). The images are aligned to the first active site residue. Key to images: a) matrilysin (human, M10.008), b) collagenase 1 (human, M10.001), c) gelatinase A (human, M10.003), d) gelatinase B (human, M10.004), e) membrane-type 1 matrix metalloproteinase (human, M10.014), f) serralysin (*Serratia marcescens*, M10.051)

contain more than one peptidase unit in separate polypeptide chains; these we term **complex peptidases**. We use a special type of identifier starting in "X" for the compound and complex peptidases. In addition, a conventional *MEROPS* identifier is assigned to each of the individual peptidase units. For example, the somatic form of peptidyl-dipeptidase A (angiotensin-converting enzyme) is XM02-001 (Corvol *et al*., 2004), and its two peptidase units are M02.001 and M02.004. There is a summary page in the database for XM02-001 in addition to the standard pages for M02.001 and M02.004. Other examples of compound peptidases are meprin A complex (XM12-001; (Bertenshaw *et al*., 2004)) and carboxypeptidase D (XM14-001;(Fricker, 1998)). Examples of complex peptidases are the proteasome (XT01-001; (Seemuller *et al*., 2004)), AAA endopeptidase complex (XM41-001; (Thorsness *et al*., 2004)), eukaryote signal peptidase (XS26-001; (Walker, 2004)) and the tricorn complex (XP01-001; (Tamura *et al*., 2004)). The tricorn peptidase complex is unique in that the components belong to different peptidase families. Peptidases that are homo-oligomers require no special classification in *MEROPS* because a single identifier can encompass all the peptidase units.

3.3.6. Peptidase inhibitors

The *MEROPS* database also includes the protein inhibitors of peptidases (Rawlings *et al.*, 2004). Many inhibitors bind to the peptidase in a substrate-like way, except that the complex is stable even if hydrolysis occurs. This mechanism is known as the **Laskowski mechanism** after the scientist who characterized it (though it is also known as the **standard mechanism** (Laskowski *et al*., 2000)). The residue that interacts with the nucleophile of the peptidase is known as the **reactive site** residue. An example of an inhibitor that uses the Laskowski mechanism is the turkey ovomucoid third domain (I01.003).

A second mechanism is known as a **trapping reaction**. This kind of reaction is specific for endopeptidases because it depends upon the cleavage of an internal bond in the inhibitor that triggers a conformational change which either traps the enzyme, for example in the case of alpha$_2$-macroglobulin inhibition (I39.001; (Barrett, 1981)), or disrupts the active site of the peptidase, for example alpha$_1$-peptidase inhibitor (I04.001; (Huntington *et al*., 2000)). Alpha$_2$-macroglobulin is able to inhibit a wide range of endopeptidases of every catalytic type because it contains a long loop containing bonds susceptible to proteolysis known as the **bait region**.

Generally, inhibitors are classified in a similar way to peptidases, and the classification is one of **inhibitor units**, an inhibitor unit being that segment of the sequence that contains a single reactive site (or bait region). There is a similar hierarchical classification of clan, family and inhibitor.

3.3.7. Compound inhibitor

At least 12 of the families of peptidase inhibitors contain what we term **compound inhibitors**; these are families I1, I2, I3, I8, I12, I15, I17, I19, I20, I25, I27 and I31. The compound inhibitors are proteins that contain multiple inhibitor units. The

number of inhibitor units ranges from 2-15 (Rawlings *et al.*, 2004). The identifier for each of these compound inhibitors starts with the letter "L" followed by the name of the family to which the peptidase units belong, a hyphen, and a serial number. For example, ovomucoid contains three inhibitor units (Kato *et al.*, 1987). These are I01.001, I01.002 and I01.003, and whole protein has the identifier LI01-001. The summary page for the compound inhibitor LI01-001 contains a diagram that shows how the individual units are arranged. A few compound inhibitors are known that contain units from more than one family of inhibitors. These have identifiers that start "LI90", and an example is chelonianin (LI90-003), which contains a domain related to I2 (I02.022) and a domain related to I17 (I17.004).

3.3.8. Type peptidase, type inhibitor

A *type peptidase* is nominated for each family and subfamily. All peptidases that are homologous to the type peptidase are members of this family. Similarly, a *type inhibitor* is nominated for each inhibitor family.

3.3.9. Families

The term *family* is used to describe a group of peptidases or peptidase inhibitors each of which can be proved to be homologous to the type example. The homology is shown by a significant similarity in amino acid sequence either to the type example itself or to another protein that has been shown to be homologous to the type example and thus a member of the family. The relationship must exist in the peptidase unit at least. A family can contain a single peptidase if no homologues are known, and a single gene product such as a virus polyprotein can contain more than one peptidase each assigned to a different family.

Some families are divided into *subfamilies* because there is evidence of a very ancient divergence within the family. Typically, the divergence corresponds to more than 150 accepted point mutations per 100 amino acid residues, which would represent an event 2,500 million years ago for a family with a typical evolutionary rate of 0.6 substitutions per amino acid site per 1,000 million years. A putative protein sequence that is very divergent from known peptidases in the family does not normally found a new subfamily but is described as "unassigned" until more is known about it.

At the time of writing, there are nearly 200 families of peptidases in *MEROPS* (Release 7.2). The naming of the families follows the system introduced by Rawlings and Barrett (Rawlings *et al.*, 1993) in which each family is named with a letter denoting the catalytic type (S, C, T, A, G, M or U, for serine, cysteine, threonine, aspartic, glutamic, metallo- or unknown), followed by an arbitrarily assigned number. For example, the caspase family of cysteine peptidases is C14. When a family disappears, usually because it is merged with another, the family name is not re-used. For this reason, there are interruptions in the numerical sequences of families that are of no current significance.

MEROPS (Release 7.2) contains 52 families of peptidase inhibitors. Because a number of families contain inhibitors of peptidases of more than one catalytic type,

it is not feasible to name the families of inhibitors according to catalytic types of the peptidases inhibited. Consequently a single series of inhibitor family names is used, formed from the letter "I" followed by a serial number. For example, family I4 (the "serpins") contains serine peptidase inhibitors such as alpha$_1$-antichymotrypsin (I04.002), but also the viral serpin CrmA (I04.028), which additionally inhibits the cysteine peptidase caspase 1 (Komiyama *et al.*, 1994).

3.3.10. Clans

In a *clan* we include all the modern-day peptidases that we believe to have arisen from a single evolutionary origin of peptidases, although they commonly have diverged so far that they now belong in more than one family. The homology of peptidases in different families in a clan is most clearly shown by their similar protein folds. The significance of the similarity can often be quantified by use of the DALI program (Holm *et al.*, 1997). When structures are not available, the order of catalytic-site residues in the polypeptide chain and sequence motifs around them may provide less direct evidence of homology at the clan level. Each clan is identified with two letters the first of which represents the catalytic type of the families included in the clan. The letter "P" is used for a clan containing families of more than one of the catalytic types serine, threonine and cysteine. Some families cannot yet be assigned to any clan, and when a formal assignment is required, such a family is described as belonging to clan A-, C-, M-, S-, T- or U-, according to the catalytic type. Some clans are divided into subclans because there is evidence of a very ancient divergence within the clan. Clan MA contains subclan MA(E), the gluzincins, and subclan MA(M), the metzincins. Clan PA is divided into subclan PA(S), containing families of serine peptidases, and subclan PA(C), containing families of cysteine peptidases. About 50 clans of peptidases are recognised in *MEROPS* (Release 7.2).

The families of proteins that inhibit peptidases are assigned to clans in similar ways to the families of peptidases. *MEROPS* (Release 7.2) contains 32 clans of inhibitors. Identifiers are taken from the ranges IA-IZ and JA-JZ.

3.3.11. Strengths of the MEROPS classification system

Peptidases and their inhibitors represent a hot-spot of scientific research on which thousands of scientists are working worldwide in academia and industry. The *MEROPS* database provides the community with a comprehensive, integrated resource. The hierarchical system of classification of peptidases and inhibitors that *MEROPS* provides is now accepted generally as authoritative. The *MEROPS* system allows for the efficient storage and retrieval of information – both within the database itself and beyond.

REFERENCES

Auld, D. (2004) Catalytic mechanisms of metallopeptidases. *In* Handbook of Proteolytic Enzymes (A. J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 268–289, Elsevier, London.

Auld, D.S. (2004) Carboxypeptidase A. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 812–821, Elsevier, London.

Barrett, A.J. (1981) alpha$_2$-Macroglobulin. Methods Enzymol 80, 737–754.

Barrett, A.J. and Chen, J.M. (2004) Thimet oligopeptidase. *In* "Handbook of Proteolytic Enzymes" (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 352–356, Elsevier, London.

Bauer, K. (2004) Pyroglutamyl-peptidase II. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 303–307, Elsevier, London.

Bertenshaw, G.P. and Bond, J.S. (2004) Meprin A. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 599–605, Elsevier, London.

Broder, D.H. and Miller, C.G. (2003) DapE can function as an aspartyl peptidase in the presence of Mn^{2+}. J Bacteriol *185*, 4748–4754.

Chapot-Chartier, M.P. and Mistou, M.Y. (2004) PepC aminopeptidase of lactic acid bacteria. *In* "Handbook of Proteolytic Enzymes" (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1202–1204, Elsevier, London.

Chave, K.J., Galivan, J. and Ryan, T.J. (2004) gamma-Glutamyl hydrolase. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1396–1399, Elsevier, London.

Chen, J.M. and Barrett, A.J. (2004) Dipeptidyl-peptidase III. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 809–812, Elsevier, London.

Corvol, P., Eyries, M. and Soubrier, F. (2004) Peptidyl-dipeptidase A/angiotensin I-converting enzyme. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 332–346, Elsevier, London.

Creemers, J.W.M. and Van de Ven, W.J.M. (2004) Furin. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1858–1861, Elsevier, London.

Dalbey, R.E. (2004) Signal peptidase I. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J. F. Woessner, Eds.), pp. 1981–1985, Elsevier, London.

Dando, P.M. (2004) Mammalian pyroglutamyl-peptidase I. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1389–1393, Elsevier, London.

Dudley, E.G. and Steele, J.L. (2004) Dipeptidase DA. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 2052–2053, Elsevier, London.

Fricker, L.D. (1998) Metallocarboxypeptidase D. *In* "Handbook of Proteolytic Enzymes" (A. J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1349–1351, Academic Press, London.

Graf, L., Szilagyi, L. and Venekei, I. (2004) Chymotrypsin. *In* "Handbook of Proteolytic Enzymes" (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1495–1501, Elsevier, London.

Hartley, B.S. (1960) Proteolytic enzymes. Annu Rev Biochem *29*, 45–72.

Holm, L. and Sander, C. (1-1-1997) Dali/FSSP classification of three-dimensional protein folds. Nucleic Acids Res *25*, 231–234.

Hooper, N.M. (2004a) Membrane dipeptidase. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 994–997, Elsevier, London.

Hooper, N.M. (2004b) X-Trp aminopeptidase. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1013–1014, Elsevier, London.

Huntington, J.A., Read, R.J. and Carrell, R.W. (19-10-2000) Structure of a serpin-protease complex shows inhibition by deformation. Nature *407*, 923–926.

Isaya, G. (2004) Mitochondrial intermediate peptidase. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 366–369, Elsevier, London.

James, M.N.G. (2004) Catalytic pathway of aspartic peptidases. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 12–19, Elsevier, London.

Kapust, R.B. and Waugh, D.S. (2000) Controlled intracellular processing of fusion proteins by TEV protease. Protein Expr.Purif. *19*, 312–318.

Kataoka, Y., Takada, K., Oyama, H., Tsunemi, M., James, M.N. and Oda, K. (6-6-2005) Catalytic residues and substrate specificity of scytalidoglutamic peptidase, the first member of the eqolisin in family (G1) of peptidases. FEBS Lett *579*, 2991–2994.

Kato, I., Schrode, J., Kohr, W.J. and Laskowski, M.J. (13-1-1987) Chicken ovomucoid: determination of its amino acid sequence, determination of the trypsin reactive site, and preparation of all three of its domains. Biochemistry *26*, 193–201.

Komiyama, T., Ray, C.A., Pickup, D.J., Howard, A.D., Thornberry, N.A., Peterson, E.P. and Salvesen, G. (29-7-1994) Inhibition of interleukin-1 beta converting enzyme by the cowpox virus serpin CrmA. An example of cross-class inhibition. J Biol Chem *269*, 19331–19337.

Laskowski, M. and Qasim, M. A. (7-3-2000) What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? Biochim Biophys Acta *1477*, 324–337.

Menard, R. and Storer, A.C. (2004a) Papain. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1125–1128, Elsevier, London.

Menard, R. and Sulea, T. (2004b) Cathepsin X. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1113–1116, Elsevier, London.

Miller, C.G. and Broder, D.H. (2004) Peptidase T. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 946–948, Elsevier, London.

Mortensen, U.H., Olesen, K. and Breddam, K. (2004) Serine carboxypeptidase C including carboxypeptidase Y. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1919–1923, Elsevier, London.

Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (1992) "Enzyme Nomenclature 1992", Academic Press, Orlando.

Polgar, L. (2004a) Catalytic mechanisms of cysteine peptidases. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1072–1079, Elsevier, London.

Polgar, L. (2004b) Catalytic mechanisms of serine and threonine peptidases. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1440–1448, Elsevier, London.

Ponting, C.P. and Pallen, M.J. (15-10-1999) beta-Propeller repeats and a PDZ domain in the tricorn protease: predicted self-compartmentalisation and C-terminal polypeptide-binding strategies of substrate selection. FEMS Microbiol Lett. *179*, 447–451.

Rao, M.B., Tanksale, A.M., Ghatge, M.S. and Deshpande, V.V. (1998) Molecular and biotechnological aspects of microbial proteases. Microbiol Mol Biol Rev. *62*, 597–635.

Rawlings, N.D. and Barrett, A.J. (1990) Bone morphogenetic protein 1 is homologous in part with calcium-dependent serine proteinase. Biochem J *266*, 622–624.

Rawlings, N.D. and Barrett, A.J. (1993) Evolutionary families of peptidases. Biochem J *290*, 205–218.

Rawlings, N.D., Tolle, D.P. and Barrett, A.J. (5-1-2004) Evolutionary families of peptidase inhibitors. Biochem J *378*, 705–716.

Sadler, J.E. (2004) Enteropeptidase. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1513–1517, Elsevier, London.

Schechter, I. and Berger, A. (1967) On the size of the active site in proteases. I. Papain. Biochem Biophys Res Commun *27*, 157–162.

Seemuller, E., Dolenc, I. and Lupas, A. (2004) Eukaryotic 20S proteasome. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 2068–2077, Elsevier, London.

Sohar, I., Sleat, D. E. and Lobel, P. (2004) Tripeptidyl-peptidase I. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1893–1896, Elsevier, London.

Southan, C. (2001) A genomic perspective on human proteases as drug targets. Drug Discov.Today *6*, 681–688.

Suzuki, F., Murakami, K., Nakamura, Y. and Inagami, T. (2004) Renin. *In* "Handbook of Proteolytic Enzymes" (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 54–61, Elsevier, London.

Tamura, T. and Baumeister, W. (2004) Tricorn protease. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 2031–2033, Elsevier, London.

Tang, J. (2004) Pepsin A. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 19–28, Elsevier, London.

Thorsness, M.K. and Thorsness, P.E. (2004) Mitochondrial *m*- and *i*-AAA proteases. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 798–801, Elsevier, London.

Tomkinson, B. (2004) Tripeptidyl-peptidase II. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1882–1885, Elsevier, London.

Turk, B., Turk, D., Dolenc, I., Turk, D. and Turk, V. (2004) Dipeptidyl-peptidase I. *In* "Handbook of Proteolytic Enzymes" (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1192–1196, Elsevier, London.

Turner, A.J. (2004) Membrane alanyl aminopeptidase. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 289–294, Elsevier, London.

Vasilyeva, O., Kolygo, K., Leonova, Y., Potapenko, N. and Ovchinnikova, T. (28-8-2002) Domain structure and ATP-induced conformational changes in *Escherichia coli* protease Lon revealed by limited proteolysis and autolysis. FEBS Lett. *526*, 66.

Walker, S. J. and Lively, M. O. (2004) Signal peptidase (eukaryote). *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1991–1995, Elsevier, London.

Wilkinson, K.D. (2004) Ubiquitin C-terminal hydrolase. *In* Handbook of Proteolytic Enzymes (A.J. Barrett, N.D. Rawlings, and J.F. Woessner, Eds.), pp. 1225–1229, Elsevier, London.

CHAPTER 11

CYSTEINE PROTEASES

ZBIGNIEW GRZONKA, FRANCISZEK KASPRZYKOWSKI
AND WIESŁAW WICZK*

Faculty of Chemistry, University of Gdańsk, Poland
**ww@chem.univ.gda.pl*

1. INTRODUCTION

Cysteine proteases (CPs) are present in all living organisms. More than twenty families of cysteine proteases have been described (Barrett, 1994) many of which (*e.g.* papain, bromelain, ficain , animal cathepsins) are of industrial importance. Recently, cysteine proteases, in particular lysosomal cathepsins, have attracted the interest of the pharmaceutical industry (Leung-Toung *et al*, 2002). Cathepsins are promising drug targets for many diseases such as osteoporosis, rheumatoid arthritis, arteriosclerosis, cancer, and inflammatory and autoimmune diseases. Caspases, another group of CPs, are important elements of the apoptotic machinery that regulates programmed cell death (Denault and Salvesen, 2002). Comprehensive information on CPs can be found in many excellent books and reviews (Barrett *et al*, 1998; Bordusa, 2002; Drauz and Waldmann, 2002; Lecaille *et al.*, 2002; McGrath, 1999; Otto and Schirmeister, 1997).

2. STRUCTURE AND FUNCTION

2.1. Classification and Evolution

Cysteine proteases (EC.3.4.22) are proteins of molecular mass about 21-30 kDa. They catalyse the hydrolysis of peptide, amide, ester, thiol ester and thiono ester bonds. The CP family can be subdivided into exopeptidases (*e.g.* cathepsin X, carboxypeptidase B) and endopeptidases (papain, bromelain, ficain, cathepsins). Exopeptidases cleave the peptide bond proximal to the amino or carboxy termini of the substrate, whereas endopeptidases cleave peptide bonds distant from the N- or C-termini. Cysteine proteases are divided into five clans: CA (papain-like enzymes),

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 181-195

CB (viral chymotrypsin-like CPs), CC (papain-like endopeptidases of RNA viruses), CD (legumain-type caspases) and CE (containing His, Glu/Asp, Gln, Cys residues in the catalytic cleft) (Barrett, 1994, 1998; Rawlings *et al.*, this volume). The majority of CPs that have been characterized are evolutionarily related to papain and share a common fold. They are synthesized as inactive precursors with a N-terminal propeptide and a signal peptide. Some peptidases of family C1 have C-terminal extensions. Activation requires proteolytic cleavage of the N-terminal proregion that also functions as an inhibitor of the enzyme. Most CPs are inhibited by E-64, cystatins and many synthetic inhibitors (Otto and Schirmeister, 1997; Grzonka *et al.*, 2001).

2.2. Papain

Papain (EC 3.4.22.2) is the best known cysteine protease. It was isolated in 1879 from the fruits of *Carica papaya* and was also the first protease for which a crystallographic structure was determined (Drenth *et al.*, 1968; Kamphuis *et al.*, 1984). The crude dried latex of papaya fruit contains a mixture of at least four cysteine proteases (papain, chymopapain, caricain, glycyl endopeptidase) and other enzymes (Baines and Brocklehurst, 1979). Crude papain of the highest quality and activity is found in sunny regions of constant humidity throughout the year. Methods of purification of papain include water extraction with reducing and chelating agents, salt precipitation and solvent extraction. Very pure papain is obtained by affinity chromatography methods. Papain is composed of 212 amino acids with three internal disulphide bridges, resulting in a molecular weight of 23.4 kDa. It is relatively basic protein, with a pI of 8.75. Its three-dimensional structure reveals that the enzyme is composed of two domains of similar size with the active cleft located between them (Fig. 1).

The general mechanism of cysteine protease action has been very well studied, with papain as the model enzyme. The enzymatic activity of papain is exerted by a catalytic dyad formed by Cys^{25} and His^{159} residues, which in the pH interval 3.5-8.0 form an ion-pair (Fig. 2). Asn^{175} is important for orientation of the imidazolium ring of the histidine in the catalytic cleft. The reactive thiol group of the enzyme has to be in the reduced form for catalytic activity. Thus, the cysteine proteases require a rather reducing and acidic environment to be active. The formation of an intermediate, S-acyl enzyme moiety, is a fundamental step in hydrolysis. This intermediate is formed *via* nucleophilic attack of the thiolate group of the cysteine residue on the carbonyl group of the hydrolysed amide (ester) bond with the release of the C-terminal fragment of the cleaved product. In the next step, a water molecule reacts with the intermediate, the N-terminal fragment is released, and the regenerated free CP molecule can begin a new catalytic cycle (Storer and Menard, 1994).

The active site residues Cys^{25} and His^{159} are positioned on opposite sides of the cleft. A number of structures of papain complexes with ligands and inhibitors have been elucidated by X-ray crystallography. Following the notation of Schechter and Berger (1967), the substrate pocket of papain binds at least seven amino acid

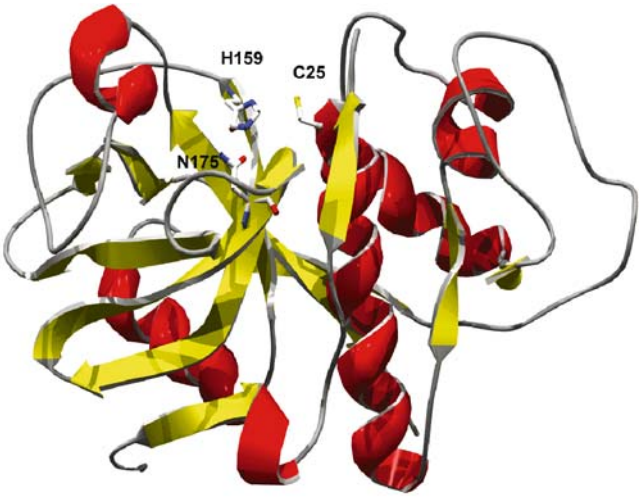


Figure 1. Ribbon representation of the three-dimensional structure of papain (Kamphuis *et al*., 1984)

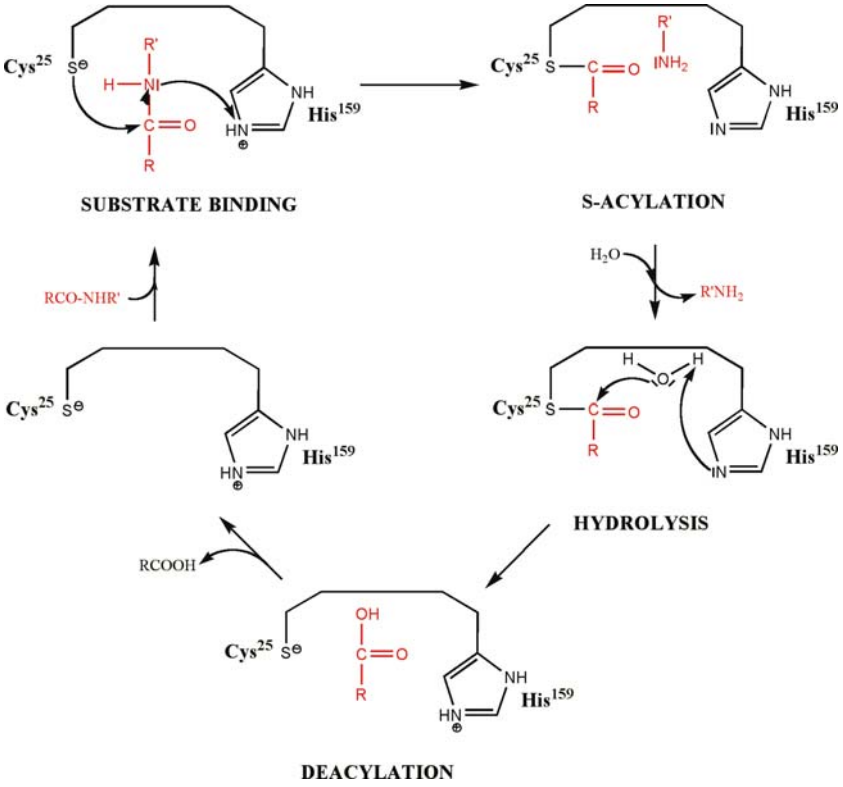


Figure 2. Enzymatic mechanism of protein hydrolysis by cysteine proteases

residues in appropriate S_n and S_n' subsites (Fig. 3). On the basis of kinetic and structural data Turk *et al.* (1998) proposed that only five subsites are important for substrate binding. According to their proposal, the S_2, S_1 and S_1' subsites are important for both backbone and side-chain binding, whereas the S_3 and S_2' pockets are crucial only for amino acid side-chain binding. A preference for those substrates containing a bulky hydrophobic chain (Phe, Leu, Ile *etc.*) in P_2 position was found; the amino acid residue in position P_1 of the substrate influences substrate binding to the enzyme to a lesser degree. There is some preference for basic amino acids (Arg, Lys) in this position but Val is not accepted. The S_3 binding site of the enzyme is less constrained; it can accommodate different amino acids side chains. Generally, papain possesses fairly broad specificity and can cleave various peptide bonds. The optimal activity of papain occurs at pH 5.8–7.0 and at temperature 50–57° C when casein is used as the substrate. Papain is stable and active for several months when stored at 4° C. Decreased activity during storage is due to oxidation of the active site thiol group. This oxidation can be partially reversed by thiol reagents (cysteine, mercaptoethanol, dimercaptopropanol *etc.*).

2.3. Bromelain

The name 'bromelain' was originally given to the mixture of proteases found in the juice of the stem and fruit of pineapple (*Ananas comosus)*. Even now, bromelain is still used as the collective name for enzymes found in various members of the Bromeliaceae family. The major endopeptidase present in extracts of plant stem is termed 'stem bromelain', whereas the major enzyme fraction found in the juice of the pineapple fruit is named 'fruit bromelain'. Some other minor cysteine endopeptidases (ananain, comosain) are also found in the pineapple stem.

Stem bromelain (EC 3.4.22.32) belongs to the papain family. It is a glycosylated single-chain protein of molecular weight 24.5 kDa. It contains 212 amino acid residues, including seven cysteines, one of which is involved in catalysis. The other six are associated in pairs forming three disulphide bridges. The crystal structure of stem bromelain has not yet been reported. Stem bromelain can be purified

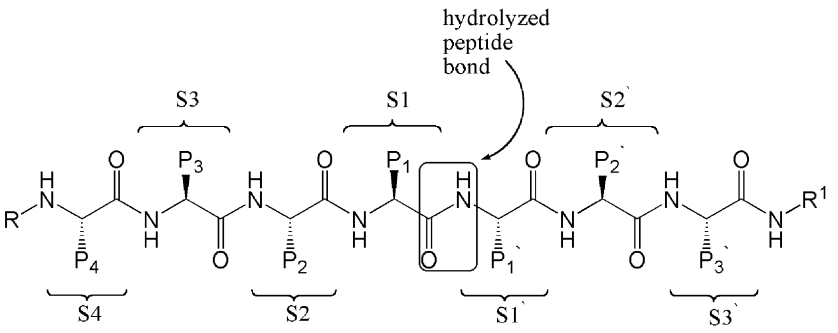


Figure 3. Interaction of papain with substrate

from dried pineapple stem powder by cation-exchange or affinity chromatography methods (Rowan *et al*, 1990). Pure stem bromelain is stable when stored at −20° C. The pH optimum for bromelain activity is 6–8.5 for most of its substrates, and the temperature optimum range of this enzyme is 50 to 60° C. Cysteine is commonly used as an activating compound for bromelain, other thiols being less effective. Stem bromelain has high proteolytic activity for protein substrates, with a preference for polar amino acids in the P_1 and P_1' positions. It has strong preference for Z-Arg-Arg-NHMec among small molecule substrates. It is scarcely inhibited by chicken cystatin and very slowly inactivated by E-64.

Fruit bromelain (EC 3.4.22.33), the major endopeptidase present in the juice of the pineapple fruit, is immunologically distinct from stem bromelain. Fruit bromelain is a single-chain glycosylated protein of molecular weight 25 kDa. It has much higher proteolytic activity compared to stem bromelain and a broader specificity for peptide bonds.

2.4. Ficain (ficin)

Ficain (EC 3.4.22.3; synonym: ficin) is the name for the cysteine protease isolated from dried latex of *Ficus glabrata*. It is also present in other species of *Ficus, e.g. F. carica, F. elastica.* Ficain can be purified by gel filtration followed by covalent chromatography (Paul *et al*, 1976). The optimum pH range is from 5 to 8, whereas the temperature optimum is from 45 to 55° C. Ficain requires cysteine or other reducing agents for activation. The enzyme has broad specificity with the acceptance of hydrophobic amino acid residues (Phe, Leu, Val) in the S_2 pocket. Ficain like papain is inhibited by chicken cystatin.

2.5. Cathepsins

Lysosomal cathepsins are an important group of enzymes that are responsible for a number of physiological processes including cellular protein degradation (Brömme and Kaleta, 2002). All cathepsins have mature domains of 214–260 amino acids. The structure of cathepsins shows an L-domain containing the active cysteine residue and a conserved α-helix and R-domain with the histidine residue and four to six β-strands. With the exception of cathepsin S, human cathepsins have acidic pH optima characteristic of the lysosomal compartment, and they are rapidly inactivated at neutral pH. Cathepsins have different specificities which are related to their specific functions in different tissues (Lecaille *et al*, 2002).

3. INDUSTRIAL APPLICATIONS OF CYSTEINE PROTEASES

Proteases, which firmly maintain first place in the world enzyme market, play an important role in biotechnology. The cysteine proteases of plants and animal cathepsins are of considerable commercial importance due to their strong proteolytic activity against a broad range of protein substrates. Most industrial applications of these enzymes are described in excellent books and review articles published

Table 1. Major industrial applications of cysteine proteases

Application	Enzymes used	Reason (uses)
Biological detergent	papain, bromelain	protein stain removing
Baking industry	bromelain, papain	lowering the protein level of flour in biscuit manufacturing, dough relaxation, preventing dough shrinkback, better bread volume, crumbliness and browning uniformity
Brewing industry	bromelain, papain	removing cloudines during storage of beers, spliting proteins in the malt
Dairy industry	bromelain, papain	whey hydrolyzates, sweetener, cheese rippening
Photographic industry	ficin	dissolving gelatin of the scraped film allowing to recovery of silver present
Food industry	bromelain, papain, cathepsins	tenderizer for meat, make high-level nutriments, make soluble protein products and breakfast, cereal and beverage, gelatin stabilization, health food, dry fermented food rippening
Waste removing (effluent)	bromelain, papain	lowering viscosity of water extract (stick water), protein and peptides production
Chitooligosaccharides production	crude bromelain, crude papain	chitosan depolymerization to use in pharmacy, animal food, medicine
Sea food	bromelain, papain	surimi production, protein hydrolyzates
Cosmetic industry	bromelain, papain	peeling effect, tooth whitening, can help to dispel taches ad pimples, clean face
Parmaceutic industry and medicine	bromelain, papain	kill the lymphatic leukemia cells, probacteria, parasite and bacillus tuberculars, helping diminish inflammation, normalize the functioning of the gallbladder, alleviating pain and promote digestion, soft lens cleaning
Textile	bromelain, papain	used for processing wool, boiling off cocoons and refining silks
Leather industry	papain	depilatory for tanning the leathers
Forage (animal‘s food)	bromelain, papain	to increase availability and inversion of proteins decreasing the cost of forages and exploiting sources of protein
Chemical industry (organic sythesis)	bromelain, papain	synthesis of aspartam, antitumor compounds, bioactive peptides

in recent years (Adler-Nissen, 1986; Vilhelmsson, 1997; Godfrey and West, 1996; Uhlig, 1998; Rao *et al*, 1998; Leisola *et al*, 2001; Shahidi and Kamil, 2001; Sentandreu *et al*, 2002; Clemente, 2000; Aehle, 2004; Liu *et al*, 2004). In Table 1 some major industrial applications are presented.

3.1. Beer and Alcohol Production

Light and clear beers are preferred by consumers. Different ingredients used during beer manufacture incorporate proteins which form insoluble complexes that appear

as a permanent haze. When the beer is chilled the insolubility increases and a more intense haze, known as chill-haze, is produced. Treatment with a proteolytic enzyme (usually crude papain or bromelain) results in a beer that remains clear and bright when chilled. Enzyme serum is also excellent as a wort clarifier (Esnault, 1995; Jones, 2005). Currently papain is not so widely used because of the trend for additive free beers prevailing in some European countries.

3.2. Baking Industry

Proteases are used in the baking industry because dough may be prepared more quickly if the gluten it contains has been partially hydrolysed. When high-gluten varieties of wheat are used the gluten must be extensively degraded for making biscuits or preventing shrinkage of commercial pie pastry. Bromelain has been widely used in the baking industry because of its rapid rate of reaction, broad pH and temperature optima and its lack of amylase or pentosanase side activities. Protease treatment improves dough relaxing and bread volume, prevents dough shrink back, and allows faster bakery throughput (Tanabe *et al*., 1996).

3.3. Food Processing

Hydrolysis of animal or vegetable food proteins is carried out for different purposes: to improve nutritional characteristics, to retard deterioration, the modification of different functional properties (solubility, foaming, coagulation, and emulsifying capacities), the prevention of undesired interactions, to change flavours and odours, and the removal of toxic or inhibitory factors, among others. Enzymatic hydrolysis is strongly preferred over chemical methods because it yields hydrolysates containing well-defined peptide mixtures and avoids the destruction of L-amino acids and the formation of toxic substances. Cysteine proteases, especially papain and bromelain, are widely used to prepare protein hydrolysates having excellent taste properties because of the absence of bitterness. Seafood (Vilhelmsson, 1997; Aspmo *et al*., 2005), eggs (Lee and Chen, 2002) and vegetable (soya, wheat, rice, sunflower, sesame and maize - Wu *et al*., 1998; Bandyopadhyay and Ghosh, 2002) protein hydrolysates not only provide excellent enhanced flavour in a wide range of foods but also improve protein assimilation (Adler-Nissen, 1986; Clemente, 2000).

Caseins and whey are some of the important protein substrates available in nature. Whey proteins generate a significant increase in foam formation and stable foam structure that can be reduced by proteolysis (Lieske and Konrad, 1996). Hydrolysis of milk proteins reduce the allergenic properties of dairy products. Milk protein hydrolysates are also used in health and fortifying sports drinks, in infant and low-digestible enteral nutrition and dietetic food.

Proteinases are widely applied in the formulation of marinades and tenderising recipes. Softness and tenderness have been identified as the most important factors affecting consumer satisfaction and the perception of taste. Tenderisation can be effected by breaking the cross-links between the fibrous protein of meat (collagen

and elastin) or by breaking meat into shreds. The traditional enzymes for this are papain, bromelain or ficin (Godfrey and West, 1996) which are sprayed or dusted onto meat. However, native meat enzymes – cathepsins and calpains – play a special role in tenderising meat by controlled ageing (Sentandreu *et al.*, 2002; Thomas *et al.*, 2004). Meat from older animals remains tough but can be tenderised by injecting inactive papain into the jugular vein of the live animal shortly before slaughtering. Upon slaughter, the resultant reducing conditions cause the accumulation of free thiols in the muscle, activating the papain and hence tenderising the meat. This is a very effective process as only 2–5 ppm of inactive enzyme need to be injected. Recently, however, it has been found that this destroys the animal's heart, liver and kidneys which cannot be sold. Papain activity is difficult to control and persists into the cooking process. Papain and bromelain as well as endogenous cysteine proteases are used for accelerated ripening of dry fermented sausages (Diaz *et al.*, 1996) and dry-cured ham (Scannell *et al.*, 2004). The activity of endogenous muscle cysteine proteases (mainly cathepsins) activated during cooking caused myosin degradation and subsequent loss of texture. In surimi production, too much cysteine protease activity is also undesirable (An *et al.*, 1996), therefore proteinase inhibitors (Gracia-Carreño, 1996) are applied to prevent gel weakening (Kang and Lanier, 2000; Rawdkuen *et al.*, 2004). Other applications include: producing dehydrated beans, baby food, food that can be easily digested by the patients, soft sweets, food deodorization (Schmidl *et al.*, 1994; Clemente, 2000).

3.4. Animal Feed

The addition of papain to some mixed forages can greatly increase the availability of protein, decreasing the cost of the forage and exploiting sources of protein (Wong *et al.*, 1996). An important application of proteases in the pet food industry is to produce a digest which liquefies the raw material and creates an acceptable flavour. This is then coated onto or mixed into dry pet food to improve its palatability.

3.5. By-product Utilization

Recently, chitosan-related materials have received a considerable amount of attention because they are useful in the food (Muzzarelli, 1996) and agriculture (Koga, 1999) industries and have various biological activities of interest (Ravi Kumar *et al.*, 2004). Chitosan is a deacylated derivative of chitin which is an abundant natural polysaccharide found in the exoskeleton of creatures such as crustaceans and insects, and in fungi. Chitinous material is obtained from the marine products' industry as a solid waste product. Chitosan depolymerisation enhances its water solubility and reduces solution viscosity as well as suppressing gel formation during storage. Therefore the depolymerisation of chitosan could facilitate the application of chitosan-related materials in a variety of fields. Commercial crude papain, bromelain and ficin are widely used for chitosan depolymerisation (Li *et al.*, 2005; Chang *et al.*, 2005). However, the hydrolysis of chitin and chitosan by means of

stem bromelain was the result of chitinase and chitosanase activities present in the crude enzyme and not bromelain itself (Hung *et al*., 2002).

Plant cysteine proteases are also used to improve the recovery of protein from slaughterhouse waste (Gómez-Juárez *et al*., 1999) and soy processing (Moure *et al.*, 2005). The recovered proteins are subsequently used in both the feed and food industries owing to their good nutritional value and excellent functional properties (Silva *et al*., 2002). Nowadays papain and alkaline bacterial proteases are also employed for solubilizing fish wastes (Gildberg *et al*., 2002; Guerard *et al.*, 2002) and to lower the viscosity of expressed fish fluids (stick water) in fodder manufacture, as well as to extract carotenoproteins from brown shrimps (Chakrabarti, 2002). Cysteine proteases are also used in skeletal muscle wasting (bone cleaning) and meat recovery processes. To recover this material, bones are mashed and incubated at 60° C with neutral or alkaline proteases for up to 4 hours. The meat slurry produced is used in canned meat and soups and protein-free bones are used as a source of gelatin.

Photographic films and plates essentially consist of an emulsion on a firm support of cellulose acetate, or polyester, or glass. The emulsion is composed of a suspension of minute silver halide crystals in gelatin. Spent films which have lost their usefulness could be utilized as a source of valuable chemicals recovered by means of the proteolytic action of papain (*i.e.* recovery of silver). Papain and bromelain are also applied to biodegrade polymers (Dupret *et al*., 2000; Howard, 2002; Chiellini *et al*., 2003).

3.6. Leather Industry

The bating of leather is a technique which takes place before tanning, and is employed to provide hides and skins with the requisite malleability and softness. Bating materials, which contain proteases, serve this purpose by breaking down the proteinaceous material of skins and hides. However, the proteolytic action should only be allowed to continue to a specific level to avoid destruction of the basic structure of the leather. In addition, papain also acts as a dehairing agent. A conventional dehairing process with sodium sulphide and lime is a major source of the pollution associated with the tanning industry. Several enzymatic (including protease and amylase activities) and non-enzymatic dehairing methods have evolved during the last century. Papain together with soluble silicates (water glass) can be used as a depilatory for tanning leathers, making the products smooth and shiny and eliminating the formation of chrome bearing leather waste (Saravanabhavan *et al.*, 2005).

3.7. Textile Industry

Papain can be used for processing wool, boiling off cocoons and refining silks (Freddi *et al*., 2003). As a result, the products will not shrink and will be quite soft. Natural silk and the engulfing gums produced by silk worms are both proteinaceous

in nature. Since papain can dissolve sericin but is unable to affect silk fibre protein it can be used for the refinement of the mixture of bombycine and vinegar fibre. In the past, papain has been widely used to 'shrink-proof' wool. A successful method involved the partial hydrolysis of the scale tips. This method also gave wool a silky lustre and added to its value. The method was abandoned a few years ago for economic reasons.

3.8. Cosmetic Industry

Enzyme baths containing bacteria and/or enzymes are popular as treatments for giving a smooth skin. Papain can help dispel blotches and pimples, clean the face and promote blood circulation making the skin healthier and tender. Papain and bromelain are used in face-care products to provide gentle peeling effects.

3.9. Organic Chemistry

Papain is used in the synthesis of amino acids (Rai and Taneja, 1998), biologically active peptides (Gill *et al*, 1996), anticancer drugs (Du, 2003) and polyaspartate (Soeda *et al*, 2003).

4. USE OF CYSTEINE PROTEASES IN PHARMACY AND MEDICINE

Due to their availability, proteases isolated from plants have a special place in these areas. A wide range of therapeutic benefits are claimed for bromelain, introduced as a therapeutic compound since 1957. Bromelain's principle activities include: the reversible inhibition of platelet aggregation (Morita *et al*, 1979), fibrinolytic activity (Maurer *et al*, 2000), anti-inflammatory action (Inoue *et al*, 1994), the modulation of cytokines and immunity (Desser *et al*, 1994; Munzig *et al*, 1995), skin debridement of burns (Rosenberg *et al*, 2004), anti-tumour activity (Batkin *et al.*, 1988), enhanced absorption of other drugs (Tinozzi and Venegoni, 1978), mucolytic properties (Hunter *et al*, 1957), a digestion aid (Knill-Jones *et al*, 1970), enhanced wound healing (Tassman *et al*, 1965) and cardiovascular and circulatory improvement (Taussig and Nieper, 1979). In addition to the cysteine protease, bromelain preparations also contains other biologically active compounds such as peroxidase, acid phosphatase, several protease inhibitors and organically bound calcium. It was found that isolation of the proteolytic fraction of bromelain leads to loss of the many beneficial effects observed *in vivo* for crude extracts (Taussig and Nieper, 1979). Results obtained from pharmaceutical and preclinical studies recommend bromelain as an orally given drug for complementary tumour therapy. The anti-metastatic activity of bromelain and its ability to inhibit metastasis-associated platelet aggregation as well as the growth and invasiveness of tumour cells is especially promising. The anti-invasive effect was found to be independent

of the proteolytic activity. (For a more comprehensive review of applications and activities of this complex of cysteine proteases see Kelly, 1996).

Another enzyme widely used in medical and para-medical practice is papain. This enzyme is used for wound debridement, the removal of necrotic tissue (Mekkes *et al.*, 1997), the external treatment of hard tissues, wart and scar tissue removal, acne treatment, depilation, skin cleansing treatments and as a component of tooth-paste. Papain is used in the preparation of tyrosine derivatives which are used for the treatment of Parkinsonism, and for the preparation of tetanus vaccines and immunoglobulin samples for intravenous injections (Brocklehurst *et al.*, 1981). Chymopapain is applied in the chemonucleolysis of damaged human intervertebral spinal discs (Watts *et al.*, 1975).

Although the toxicity of the above mentioned enzymes is rather low, exposure to the dust or aerosols of their solutions is harmful. Such exposure may induce asthma, rhinitis and allergy (Baur and Fruhmann, 1979; Flindt, 1978; Novey *et al.*, 1979). Papain is used in laboratory practice for artificial induction of emphysema (Martorana *et al.*, 1982) and osteoarthritis (Kopp *et al.*, 1983) in experimental animals. Anaphylaxis is one of the complications caused by chymopapain used in chemonucleolysis (Watts *et al.*, 1975; Ford, 1977; DiMaio, 1976). Others are subarachnoid haemorrhage (Buchman *et al.*, 1985), nerve injury (Mackinnon *et al.*, 1984) and intervertebral disk-space infections (Deeb *et al.*, 1985).

Cysteine proteases have also been recognized as critical enzymes in degenerative and autoimmune states. Lysosomal cysteine proteases of the papain family are involved in different pathological states. Deficiency of enzymatic activity of this group of enzymes was found to occur in two diseases: pycnodysostosis, a skeletal bone dysplasia caused by cathepsin K deficiency, and Pappilon-Lefevre syndrome, a periodontopathia caused by cathepsin C defficiency (Lecaille *et al.*, 2002). However, the major role of papain-like cysteine proteases in pathological states is not related to their deficiency but the overexpression of such enzymes or their activity outside their normal site of action. An understanding of the physiopathological functions of cysteine proteases will permit the design of new selective therapeutic agents.

Tumour cell invasion and metastasis are associated with the proteolytic activities of various types of proteases, including lysosomal proteases. Elevated expression of certain cathepsins and diminished levels of their inhibitors have been observed in several human cancers, including breast, gastric, glioma and prostate cancers, and especially in cases of aggressive cells (Lecaille *et al.*, 2002; Otto and Schirmeister, 1997).

Cathepsins of the papain family seem to play a critical role in rheumatoid arthritis (Taubert *et al.*, 2002) and atherosclerosis (Lecaille *et al.*, 2002; Otto and Schirmeister, 1997).

Cysteine proteases of the papain family play an important role in microbial (viral, bacterial) and parasitic infections (Tong, 2002; Han *et al.*, 2005). They are virulence factors and/or participate in tissue penetration, feeding, replication and immune evasion. The lack of redundancy of the cysteine proteases in these

organisms compared to their mammalian hosts makes them attractive targets for the development of new medically useful compounds.

Intense development of enzyme applications for food and animal feeds, the detergent and textile industries as well as in medicine mean that the current list of cysteine protease applications is incomplete. However, variability in the properties of plant enzymes which depend on weather conditions amongst others may well result in their dispacement by microbial enzymes. Genetic engineering techniques will be applicable not only to source valued enzymes in easy-to-grow micro-organisms but also to modify and tailor enzyme properties to consumer requirements.

REFERENCES

Adler-Nissen, J. (1986) Enzymatic hydrolysis of food proteins. Elsevier Applied Science Publisher, London, New York.

Aehle, W. (2004) Enzymes in industry. Production and application. Wiley-VCH, Weinheim.

An, H., Margo, Y.P. and Seymour, T.A. (1996) Role of endogenous enzymes in surimi gelation. Trends Food Sci. Tech. *7*, 321–326.

Aspmo, S.I., Horn, S.J. and Eijsink, V.G.H. (2005) Enzymatic hydrolysis of atlantic cod (*Gadus morhua* L.) viscera. Process Biochem. *40*, 1957–1966.

Baines, B.S. and Brocklehurst, K. (1979) A necessery modification to the preparation of papain from any high-quality latex of *Carica papaya* and evidence for structural integrity of the enzyme produced by traditional methods. Biochem. J. *177*, 541–548.

Bandyopadhyay, K. and Ghosh, S. (2002) Preparation and characterization of papain-modified sesame (*Sesamum indicum* L.) protein isolates. J. Agric. Food Chem. *50*, 6854–6857.

Barrett, A.J. (1994) Classification of peptidases. Methods Enzymol. *244*, 1–15.

Barrett, A.J., Rawlings, N.D. and Woessner, J.F. (1998) Handbook of Proteolytic Enzymes. Academic Press, San Diego.

Batkin, S., Taussig, S.J. and Szekerczes, J. (1988) Antimetastatic effect of bromelain with or without its proteolytic and anticoagulant activity. J. Cancer Res. Clin. Oncol. *114*, 507.

Baur, X. and Fruhmann, G. (1979) Allergic reactions, including asthma, to the pineapple protease bromelain following occupational exposure. Clin. Allergy *9*, 443–450.

Bordusa, F. (2002) Proteases in organic synthesis. Chem. Rev. *102*, 4817–4867.

Brocklehurst, K., Baines, B.S., Kierstan, M.P.J. (1981) Papain and other constituents of Carica papaya L. Top. Enzyme Ferment. Biotechnol. *5*, 262–335.

Brömme, D. and Kaleta, J. (2002) Thiol-dependent cathepsins: Pathophysiological implications and recent advances in inhibitor design. Curr. Pharmac. Design *8*, 1639–1658.

Buchman, A., Wright, R.B., Wichter, M.D., Whisler, W.W. and Bosch, A. (1985) Hemorrhagic complications after the lumbar injection of chymopapain. Neurosurgery *16*, 222–224.

Chakrabarti, R. (2002) Carotenoprotein from tropical brown shrimp shell waste by enzymatic process. Food Biotechnol. *16*, 81–90.

Chang, C.-L., Chang, Y.-M., Chang, C.-T. and Sung, H.-Y. (2005) Characterization of a chitosanase isolated from a commercial ficin preparation. J. Agric. Food Chem. *53*, 7579–7585.

Chiellini, E., Corti, A., D'Antone, S. and Solaro, R. (2003) Biodegradation of poly(vinyl alcohol) based materials. Prog. Polym. Sci. *28*, 963–1014.

Clemente, A. (2000) Enzymatic protein hydrolysates in human nutrition. Trends Food Sci. Tech. *11*, 254–262.

Deeb, Z.L., Schimel, S., Daffner, R.H., Lupetin, A.R., Hryshko, F.G. and Blakley, J.B. (1985) Intervertebral disk-space infection after chymopapain injection. Am. J. Roentgenol. *144*, 671–674.

Denault, J.B. and Salvesen, G.S. (2002) Caspases: Keys in the ignition of cell death. Chem. Rev. *102*, 4489–4999.

Desser, L., Rehberger, A. and Paukovits, W. (1994) Proteolytic enzymes and amylase induce cytokine production in human peripheral blood mononuclear cells *in vivo*. Cancer Biother. *9*, 253–263.

Diaz, O., Fernandéz, M., Gracia de Fernando, C.D., de la Hoz, L. and Ordòňez, J.A. (1996) Effect of the addition of papain on the dry fermented sausage proteolysis. J. Sci. Food Agric. *71*, 13–21.

DiMaio, V.J. (1976) Two anaphylactic deaths after chemonucleolysis. J. Forensic Sci. *21*, 187–190.

Drauz, K. and Waldmann, H. (2002) Enzyme Catalysis in Organic Synthesis, Wiley-VCH, Weinheim.

Drenth, J., Jansonius, J.N., Koekoek, R., Swen, H.M. and Wolthers, B.G. (1968) Structure of papain. Nature *218*, 929–932.

Du, W. (2003) Towards new anticancer drugs: a decade of advances in synthesis of camptothecins and related alkaloids. Tetrahedron *59*, 8649–8687.

Dupret, I., David, C. and Daro, A. (2000) Biodegradation of polyester-amides using a pure strain of micro-organisms or papain I. Model compounds. Polym. Degrad. Stabil. *67*, 497–505, 505–509.

Esnault, E. (1995) Beer stabilization with papain. Brew. Guard. *124*, 47–49.

Flindt, M.L. (1978) Respiratory hazards from papain. Lancet *1*, 430–432.

Ford, L.T. (1977) Chymopapain-past and present, future? Cli. Orthop. Relat. Res. *122*, 367–373.

Freddi, G., Mossotti, R. and Innocenti, R. (2003) Degumming of silk fabric with several proteases. J. Biotechnol. *196*, 101–112.

Gildberg, A., Arnesen J.A. and Carlehög, M. (2002) Utilization of cod backbone by biochemical fractionation. Process Biochem. *38*, 475–480.

Gill, I., López-Fandiño, S., Jorba, X. and Vulfson, E.N. (1996) Biologically active peptides and enzymatic approaches to their production. Enzyme Microb. Technol. *18*, 162–183.

Godfrey, T. and West, S. Eds. (1996) Industrial enzymes, Stocton Press, New York.

Gómez-Juárez, C., Casttelanos, R., Ponce-Noyala, T., Calderón, V. and Figueroa, J. (1999) Protein recovery from slaughterhouse wastes. Bioresource Technol. *70*, 129–133.

Gracia-Carreño, F.L. (1996) Proteinase inhibitors. Trends Food Sci. *7*, 197–204.

Grzonka, Z., Jankowska, E., Kasprzykowski, F., Kasprzykowska, R., Łankiewicz, L., Wiczk, W., Wieczerzak, E., Ciarkowski, J. Drabik, P., Janowski, R., Kozak, M., Jaskólski, M. and Grubb, A. (2001) Structural studies of cysteine proteases and their inhibitors. Acta. Bioch. Pol. *48*, 1–20.

Guerard, F., Guimas, L. and Binet A. (2002) Production of tuna waste hydrolysates by a commercial neutral protease preparation. J. Mol. Catal. B-Enzym. 19-*20*, 489–498.

Han, Y-S., Chang, G-G., Juo, Ch-G., Lee, H-J., Yeh, S-H., Hsu, J. T-S. and Chen, X. (2005) Papain-like protease 2 (PLP2) from severe acute respiratory syndrome coronavirus (SARS-CoV): expression, purification characterization and inhibition. Biochemistry *44*, 10349–10359.

Howard, G.T. (2002) Biodegradation of polyurethane: a review. Int. Biodeter. Biodegr. *49*, 245–252.

Hung, T.-H., Chang, Y.-M., Sung, H.Y. and Chang, C.-T. (2002) Purification and characterization of hydrolase with chitinase and chitosanase activity from commercial steam bromelain. J. Agric. Food Chem. *50*, 4666-4673.

Hunter, R.G., Henry, G.W. and Heinicke, R.M. (1957) The action of papain and bromelain on the uterus. Am. J. Ob. Gyn. *73*, 867–873.

Inoue, K., Motonaga, A., Dainaka, J., Nishimura, T., Hashii, H., Jamate, K., Ueda, F. and Kimura, K. (1994) Effect of etodolac on prostaglandin E2 biosynthesis, active oxygen generation and bradykinin formation. Prostaglandins Leukot. Essent. Fatty acids *51*, 457–462.

Jones, B.L. (2005) Endoproteases of barley and malt. J. Cereal Sci. *42*, 139–156.

Kamphuis, I.G., Kalk, K.H., Swarte, M.B.A. and Drenth, J. (1984) Structure of papain refined at 1.65Å resolution. J. Mol. Biol. *179*, 233–256.

Kang, I.S. and Lanier, T.C. (2000) Heat induced softening of surimi gel by proteinases. In Surimi and surimi seafood, Park, J.W. (Ed.). Marcel Dekker, New York, pp. 445–474.

Kelly, G.S. (1996) Bromelain: a literature review and discussion of its therapeutic applications. Alt. Med. Rev. *1*, 243–257.

Knill-Jones, R.P., Pearce, H., Batten, J. and Williams, R. (1970) Comparative trial of Nutrizym in chronic pancreatic insufficiency. Brit. Med. J. *4*, 21–24

Koga, D., Mitsutomi, M., Kono, M. and Matsumija, M. (1999) Biochemistry of chitinases. In Chitin and chitinases, Jolles, P. and Mazzurelli, R. A. A. (Eds.). Brikhäuser Verlag, Basel, Switzerland, vol *98*, pp. 111–123.

Kopp, S., Mejersjo, C. and Clemensson, E. (1983) Induction of osteoarthrosis in the guinea pig knee by papain. Oral Surg. Oral Med. Oral Pathol. *55*, 259–266.

Lecaille, F., Kaleta, J. and Brőmme, D. (2002) Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. Chem. Rev. *102*, 4459–4488.

Lee, W.C. and Chen T.C. (2002) Functional characteristics of egg white solids obtained from papain treated albumen. J. Food Eng. *51*, 263–266.

Leisola, M., Jokela, J., Pastinen, O., Turunen, O. and Schoemaker, H. (2001) Industrial use of enzymes, Eolas Publisher, Oxford.

Leung-Toung, R., Li, W., Tam, T.F. and Karimian, K. (2002) Thioldependent enzymes and their inhibitors: a review. Curr. Med. Chem. *9*, 979–1002.

Li, J., Du, Y., Yang, J., Feng, T., Li, A. and Chen, P. (2005) Preparation and characterization of low molecular weight chitosan and chito-oligomers by a commercial enzyme. Polym. Degrad. Stabil. *87*, 441–448.

Lieske, B. and Konrad, G. (1996) Physicochemical and functional properties of whey protein as affected by limited papain proteolysis and selective ultrafiltration. Int. Dairy J. *6*, 13–31.

Liu, Z., Weis, R. and Glieder, A. (2004) Enzymes from higher eukaryotes for industrial biocatalysis. Food Technol. Biotechnol. *42*, 237–249.

Mackinnon, S.E., Hudson, A.R., Llamas, F., Dellon, A.L., Kline, D.G. and Hunter, D.A. (1984) Peripheral nerve injury by chymopapain injection. J. Neurosurg. *61*, 1–8.

Martorana, P.A., Wusten, B., Van Even, P., Gobel, H. and Scharper, J. (1982) A six-month study of the evolution of papain-induced emphysema in the dog. Am. Rev. Respir. Dis. *126*, 898–903.

Maurer, H.R., Eckert, K., Grabowska, E. and Eschman, K. (2000) Use of bromelain proteases for inhibiting blood coagulation. Patent WO PCT/EP 98/04406.

McGrath, M.E. (1999) The lysosomal cysteine proteases. Annu. Rev. Biophys. Biomol. Struct. *28*, 181–204.

Mekkes, J.R., Le Poole, I.C., Das, P.K., Kammeyer, A. and Westerhof, W. (1997) *In vitro* tissue-digesting properties of krill enzymes compared with fibrinolysin/DNAse, papain and placebo. Int. J. Cell. Biol. *29*, 703–706.

Morita, A.H., Uchida, D.A. and Taussig, S.J. (1979) Chromatographic fractionation and characterization of the active platelet aggregation inhibitory factor from bromelain. Arch. Inter. Phar. Ther. *239*, 340–350.

Moure, A., Dominguez, H. and Parajó, C. (2005) Fractionation and enzymatic hydrolysis of soluble protein present in waste from soy processing. J. Agric. Food Chem. *53*, 7600–7608.

Munzig, E., Eckert, K., Harrach, T., Graf, H. and Maurer, H.R. (1995) Bromelain protease F9 reduces the CD44 mediated adhesion of human peripheral blood limphocytes to human umbilical vein endothelial cells. FEBS Lett *351*, 215–218.

Muzzarelli, R.A.A. (1996) Chitosan-based dietary foods, Carbohyd. Polym. *29*, 309–316.

Novey, H.S., Marchioli, L.E., Sokol, W.N. and Wells, I.D. (1979) Papain-induced asthma-physiological and immunological features. J. Allergy Clin. Immunol. *63*, 98–103.

Otto, H.-H. and Schirmeister, T. (1997) Cysteine proteases and their inhibitors. Chem. Rev. *97*, 133–171.

Paul, J., Malthouse, G. and Brocklehurst, K. (1976) Preparation of fully active ficin from *Ficus glabrata* by covalent chromatography and characterization of its active centre by using 2,2'-dipyridyl disulphide as a reactivity probe. Biochem. J. *159*, 221–234.

Rai, R. and Taneja, V. (1998) Papain catalysed hydantoin hydrolysis in the synthesis of amino acids. Biochem. Biophys. Res. Commun. *244*, 889–892.

Rao, M.B., Tanksale, A.M., Ghatge, M.S. and Deshpande, V.V. (1998) Molecular and biotechnological aspects of microbial proteases. Microbiol. Molec. Biol. Rev. *62*, 597–635.

Ravi Kumar, M.N.V., Muzzarelli, R.A.A., Muzzarelli, C., Sashiva, H. and Domb, A. (2004) Chitosan chemistry and pharmaceutical perspectives. Chem. Rev. *104*, 6017–6084.

Rawdkuen, S., Benjakul, S., Visessanguan, W. and Lanier, T. (2004) Chicken plasma protein: Proteinase inhibitory activity and its effect on surimi gel properties. Food Res. Int. *37*, 156–165.

Rosenberg, L., Lapid, O., Bogdanov–Bierezovsky, A., Glesinger, R., Krieger, Y., Silberstein, E., Sagi, A., Judkins, K. and Singer, A.J. (2004) Safety and efficacy of proteolytic enzyme for enzymatic burn debridement: a preliminary report. Burns *30*, 843–850.

Rowan, A.D., Buttle, D.J. and Barrett, A.J. (1990) The cysteine proteinases of the pine apple plant. Biochem. J. *266*, 869–875.

Saravanabhavan, S., Thanikaivelan, P., Rao, J.R. and Nair B.U. (2005) Silicate enhancement enzymatic dehairing: a new lime-sulfite-free process for cowhides. Environ. Sci. Technol. *39*, 3776–3783.

Scannell, A.G.M., Kenneally, P.M. and Arendt, E.K. (2004) Contribution of started cultures to the proteolytic process of a fermented non-dried whole muscle ham product. Int. J. Food Microbiol. *93*, 219–230.

Schechter, I. and Berger, A. (1967) On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. *27*, 157–162.

Schmidl, M.K., Taylor, S.L. and Nordlee, J.A. (1994) Use of hydrolysate-based products in special medical diets. Food Technol. *48*, 77–80.

Sentandreu, M.A., Coulis, G. and Ouali, A. (2002) Role of muscle endopeptidases and their inhibitors in meat tenderness. Trends Food Sci. Technol. *13*, 398–419.

Shahidi, F. and Kamil, Y.V.A.J. (2001) Enzymes from fish and aquatic invertebrates and their application in the food industry. Food Sci. Technol. *12*, 435–464.

Silva, J.G., Morais, H.A., Oliveira, A.L. and Silvestre, M.P.C. (2002) Addition effects of bovine blood globin and sodium caseinate on the characteristics of raw and cooked ham pate. Meat Sci. *63*, 177–184.

Soeda, Y., Toshima, K. and Natsumura, S. (2003) Sustainable enzymatic preparation of polyaspartate using bacterial protease. Biomacromolecules *4*, 196–203.

Storer, A.C. and Menard, R. (1994) Catalytic mechanism in papain family of cysteine peptidases. Methods Enzymol. *244*, 486–500.

Tanabe, S., Arai, S. and Watanabe, M. (1996) Modification of wheat flour with bromelain and baking hypoallergenic bread with added ingredients. Biosci. Biotech. Bioch. *60*, 1269–1272.

Tassman, G.C, Zafran, J.N, and Zayon, G.M. (1965) A double – blind crossover study of a plant proteolytic enzyme in oral surgery. J. Dent. Med. *20*, 51–54.

Taubert, H., Riemann, D., Kehlen, A., Meye, A., Bartel, F., John, V., Brandt, J., Bache, M., Wurl, P., Schmidt, H. and Weber, E. (2002) Expression of cathepsin B, D and L protein in juvenile idiopathic arthritis. Autoimmunity *35*, 221–224.

Taussig, S.J. and Nieper, H.A. (1979) Bromelain: its use in prevention and treatment of cardiovascular disease, present status. J IAPM *6*, 139–151.

Thomas, A.R., Gondoza, H., Hoffman, L.C., Oosthuizen, V. and Naudé, R.J. (2004) The role of the proteasome, and cathepsins B, L, H and D in ostrich meat tenderization. Meat Sci. *67*, 113–120.

Tinozzi, S. and Venegoni, A. (1978) Effect of bromelain on serum and tissue levels of amoxycyllin. Drugs Expt. Clin. Res. *4*, 39–44.

Tong, L. (2002) Viral proteases. Chem. Rev. *102*, 4609–4626.

Turk, D., Gunčar, G., Podobnik, M. and Turk, B. (1998) Revised definition of substrate sites of papain-like cysteine proteases. Biol. Chem. *379*, 137–147.

Uhlig, H. (1998) Industrial enzymes and their application. J. Wiley and Sons, New York.

Vilhelmsson, O. (1997) The state of enzyme biotechnology in the fish processing industry. Trends Food Sci. Tech. *8*, 266–271.

Watts, C., Hutchinson, G., Stern, J. and Clark, K. (1975) Comparison of intervertebral disc disease treatment by chymopapain injection and open surgery. J. Neurosurg. *42*, 397–400.

Wong, M.H., Tang, L.Y. and Kwok F.S. (1996) The use of enzyme-digested soybean residue for feeding common carp. Biomed. Environ. Sci. *9*, 418–423.

Wu, W.U., Hettiarachchy, N.S. and Qi, M. (1998) Hydrophobicity, solubility, and emulsifying properties of soy protein peptides prepared by papain modification and ultrafiltration. J. Am. Oil Chem. Soc. *75*, 8945–8950.

CHAPTER 12

SUBTILISIN

JOHN DONLON*

Department of Biochemistry, National University of Ireland, Galway, Ireland
**john.donlon@nuigalway.ie*

1. INTRODUCTION

Proteolytic enzymes (proteases) are omnipresent in nature (see Rawlings *et al.*, Chapter 10, this volume). Subtilisins are a family of serine proteases, *i.e.* they possess an essential serine residue at the active site. This serine residue is part of a catalytic triad of Aspartate, Histidine and Serine that is very similar to that of mammalian intestinal digestive enzymes, trypsin and chymotrypsin. The subtilisin family, now known as peptidase family S8, is the second largest serine protease family. There are over 200 known members of the family, with the complete amino acid sequence established for the vast majority of them (Siezen and Leunissen, 1997). Proteolytic enzymes that utilize serine in their catalytic triad are quite ubiquitous. They include a wide range of peptidase activities, such as endopeptidases, exopeptidases and oligopeptidases. Over 20 families of serine proteases have been identified and classified as members of 6 clans on the basis of structural and functional similarities. Subtilisins are to be found in archaebacteria, eubacteria, eukaryotes and viruses. The bacterial subtilisins are the subgroup of serine proteases of greater industrial significance and have been studied extensively, with regard to improving their catalytic efficiency and stabilities. As detailed later, those subtilisins produced by selected *bacilli* have found widespread applications, especially as detergent additives.

2. GENERAL PROPERTIES OF SUBTILISINS

The serine proteases have a catalytic triad of serine, aspartate and histidine in common. A specific serine residue acts as a nucleophile and anchors the acyl-enzyme intermediate during the course of the enzyme's catalytic action, with

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 197–206

aspartate as an electrophile, and histidine as a base. It is notable that the geometric orientation of the catalytic residues is similar between families, despite different protein folds and the absence of sequence homology. The linear arrangements of the catalytic residues commonly reflect clan relationships. The catalytic triad in the chymotrypsin clan is ordered his-ser-asp; but it is ordered asp-his-ser in the subtilisin clan. Interestingly, bacterial subtilisins and mammalian serine proteases are paradigms of convergent evolution having independently arrived at this very similar catalytic triad (Rawlings and Barrett, 1993). Thus, for these serine proteinases, having unrelated ancestral precursors, convergent evolution has resulted in a very similar structural arrangement to achieve a particular catalytic mechanism.

All these enzymes catalyze the hydrolysis of peptide and ester bonds through formation of an acyl-enzyme intermediate (see Perona and Craik, 1995; Polgár, 2005 and Rawlings *et al.*, Chapter 10, this volume for detailed reviews). Briefly, after formation of enzyme-substrate complex, the carbonyl carbon of the scissile bond is attacked by the active site serine, forming a tetrahedral intermediate. In subtilisins this transition state is stabilised by hydrogen bonding to the backbone of the serine 221 (the active site nucleophile) and the side chain of asparagine 155. This transition state decays as a proton is donated from the active site histidine 64 to the amine group at the cleavage site of the substrate to liberate the first product of the reaction and simultaneous formation of the covalent acyl-enzyme intermediate. The enzyme is deacylated by nucleophilic attack by water, followed by the formation of another tetrahedral intermediate that is also stabilised by hydrogen bonding to the enzyme. This decays with proton transfer to the active site histidine and release of the second peptide product. With regard to the third member of the catalytic triad, aspartate 32, there is another characteristic trait of serine proteases, *i.e.* a resonance between its carboxylate and histidine 64 mediated by a low-barrier hydrogen bond (LBHB) influencing the reactivity of histidine 64 in a manner that is generally regarded as critical for catalysis. LBHBs are such that the hydrogen atom becomes more or less equally shared between the donor and acceptor atoms. However, the criticality of this LBHB as an inherent requirement for significant rate enhancement for subtilisin has been recently called into question (Stratton *et al.*, 2001).

Most members of peptidase family S8 are endopeptidases. Most of the family are active at neutral to, generally, mildly alkaline pH. Many of them are extremely thermostable, which make them very suited to many applications. Most of them are non-specific peptidases having high turnover numbers and with a preference to cleave at the C-terminal side of hydrophobic residues. However, thermophilic subtilisins are generally less catalytically efficient. Subtilisins accept a broader range of substrates other than peptides or proteins, so they are also used for reactions involving unnatural substrates in synthetic reactions (Moree *et al.*, 1997). In this context, subtilisins are more tolerant of changes in the nucleophile than in the carboxyl group. They are inhibited by the general serine protease inhibitors, such as nerve gases (*e.g.* diisopropyl fluorophosphate) and phenylmethanesulfonyl fluoride. The tertiary structures for several of them have been determined under various conditions. An S8 protease typically consists of three layers with a 7-stranded

β-sheet sandwiched between two layers of α-helices (Fig. 1). The structural stability of these enzymes is illustrated by subtilisin Carlsberg in neat organic solvent, showing an extremely well organised molecule (Fig. 1). Another feature of this family of proteases is the presence of one or more calcium binding sites that contribute greatly to the thermal stability of many of them.

Subtilisins have another property in common with many secreted proteases, *i.e.* their biosynthesis requires participation of an N-terminal pro-domain (see Shinde and Inouye, 2000). Such domains act as intra-molecular chaperones to greatly expedite the folding rate of the mature, stable subtilisin. This, clearly, provides nature with a clever mechanism of regulating protease activation and it also provides mankind with an approach to maintaining industrially important subtilisins in extremely stable states that can be activated at will (Takagi and Takahashi, 2003; Subbian *et al*., 2005).

The advent of recombinant DNA technology has brought about a revolution in the development of new enzymes and in our understanding of the structure/function relationships of proteins, in general. The general features of structure and function relationships of subtilisins gleaned from earlier studies have been reviewed (Jarnagin and Ferrari, 1992). They noted that most single mutations in subtilisin BPN' do not cause major structural alterations. Even multiple mutations, though

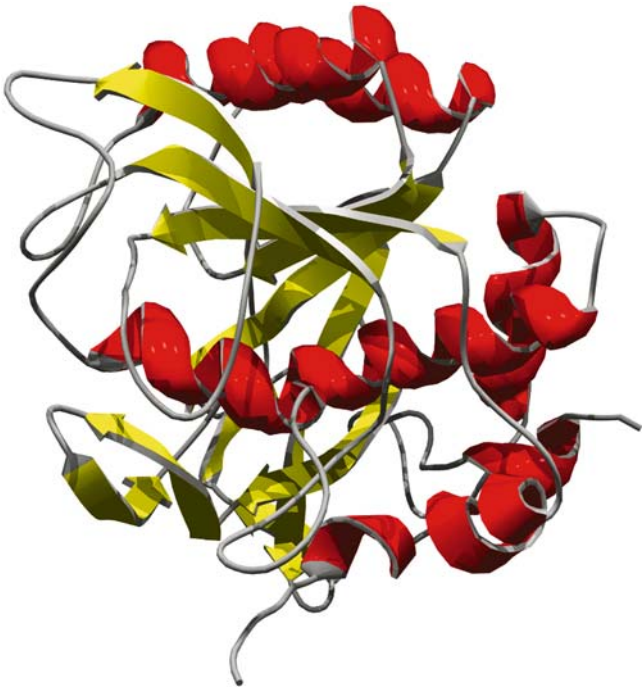


Figure 1. Subtilisin Carlsberg (E.C.3.4.21.62): Enzyme crystal structure in a neat organic solvent

they may cause local minor perturbations, do not alter overall structure to any large degree. It had been observed earlier that the subtilisin BPN' structure is very tolerant of single mutations, and this tolerance may have been necessary for survival of the enzyme during the course of evolution. This structural tolerance is not surprising if one considers that the structure of subtilisin Carlsberg is very similar to that of subtilisin BPN' while their protein sequences differ by 31%. Apparently, a significant amount of sequence variation still allows for overall structural similarities in the subtilisin family of enzymes. Though the overall structure of subtilisin is not easily perturbed by single or even multiple mutations, it is also clear that single mutations can lead to very significant affects on the catalytic efficiency, substrate preference, and stability.

3. BIOENGINEERING OF SUBTILISINS

Subtilisin has also become a paradigm for protein engineering studies. Protein engineering of subtilisin commenced in the 1960s, with a view to understanding their catalytic properties and stabilities (earlier studies comprehensively reviewed by Bryan, 2000). Since the advent of gene cloning in the early 1980s there have been many impressive studies involving genetic manipulation of subtilisins. For example, by directed evolution, subtilisin E from *Bacillus subtilis* was converted into an enzyme functionally equivalent to its thermophilic homologue, thermitase from *Thermoactinomyces vulgaris*. Thermitase, also a member of the subtilisin family, has 47% sequence homology to subtilisin BPN' (Gros *et al*, 1989). Five generations of random mutagenesis, recombination and screening created subtilisin E 5-3H5 (Zhao and Arnold, 1999). The optimum temperature of the evolved enzyme was 17°C higher and its half-life at 65°C was more than 200-fold that of wild type subtilisin E. In addition, 5-3H5 was more active towards the hydrolysis of a synthetic substrate, succinyl-Ala-Ala-Pro-Phe-p-nitroanilide, than wild type at all temperatures from 10 to 90°C. Surprisingly, even though the sequence of thermitase differs from that of subtilisin E at 157 positions, only eight amino acid substitutions were required to convert subtilisin E into an enzyme with similar thermostability. The eight substitutions, which included previously recognised stabilizing mutations (*e.g.* asparagine replacing serine at position 218 and aspartate for asparagine at residue 76), were found distributed over the surface of the enzyme. Impressively, these experiments showed that directed evolution provides a powerful tool to unveil mechanisms of thermal adaptation and that it is an effective and efficient approach to manipulating thermostability without compromising enzyme activity.

A more recent study on the stabilizing mutations in subtilisin BPN' has also greatly aided understanding of the structural basis of the thermostability of this enzyme (Almog *et al*, 2002). The rationale for this study was based on a requirement to overcome the loss of calcium due to the presence of water softeners (chelators) encountered during use of detergents (*vide infra*). Two new variants of calcium-independent subtilisin were created, where the high affinity calcium site was deleted, and then selected for increased thermostability from a panel of random mutants.

The molecular structures of these two enzymes have been compared with previously solved structures of subtilisin. Despite the variations in sequence, *etc*., the overall structures are similar but not in the N-terminal region adjacent to the deletion. One of the variants formed a disulfide bond between the new cysteine residues. This disulfide bond anchors the N- terminus and contributes to the dramatic increase in thermostability. In addition to the new disulfide bond, other mutations combined to increase its thermostability 1200-fold under chelating conditions, essentially due to stabilization of the N-terminus. More recent site directed mutagenesis have vastly improved the enzymatic half-life of calcium-free subtilisin BPN', also with potential usefulness for biotechnological applications (Strausberg *et al*., 2005).

Enzymes isolated from psychrophilic organisms (native to cold environments) generally exhibit higher catalytic efficiency at low temperatures and greater thermosensitivity than their moderate mesophilic counterparts. In an effort to understand the evolutionary process and the molecular basis of cold adaptation, directed evolution has also been employed to convert a mesophilic subtilisin-like protease from *Bacillus sphaericus*, SSII, into its psychrophilic counterpart. A single round of random mutagenesis followed by recombination of improved variants yielded a mutant with a turnover number (k_{cat}), at 10°C, increased 6.6-fold and a catalytic efficiency (k_{cat}/K_m) 9.6 times that of wild type. Its half-life at 70°C was found to be 3.3 times less than wild type. It has been noted that although there is a trend toward decreasing stability during the progression from mesophilic to psychrophilic enzymes, there is no strict correlation between decreasing stability and increasing low temperature activity. Mesophilic subtilisin, SSII, shares 77.4% sequence identity with the naturally psychrophilic protease subtilisin, S41. Although, these two subtilisins differ at 85 positions, yet just four amino acid substitutions were sufficient to generate an SSII subtilisin whose low temperature activity is greater than that of S41 (Wintrode *et al*., 2000).

The thermostability and activity of the psychrophilic protease subtilisin S41, from the Antarctic *Bacillus* TA41, was also investigated with the goal of understanding the mechanisms by which this enzyme can adapt to different selection pressures. Mutant libraries were screened to identify enzymes that acquired greater thermostability without sacrificing low-temperature activity. The half-life of a seven-amino acid substitution variant, 3-2G7, at 60°C was approximately 500 times that of wild type and far surpassed those of homologous mesophilic subtilisins. The temperature optimum of the activity of 3-2G7 was shifted upward by approximately 10 degrees C. Unlike natural thermophilic enzymes the activity of 3-2G7 at low temperatures was not compromised. The catalytic efficiency was enhanced approximately 3-fold over a wide temperature range (10 to 60°C). The activation energy for catalysis was nearly identical to wild type and close to half that of its highly similar mesophilic homologue, subtilisin SSII, indicating that the evolved S41 enzyme retained its psychrophilic character in spite of its dramatically increased thermostability. These results clearly demonstrated that it is possible to increase activity at low temperatures and stability at high temperatures simultaneously. As has been speculated, the fact that enzymes displaying both properties are not found in nature

Table 1. Commercial subtilisins used in detergents

Trade name (and producer)	Origin	T/PE[a]	Production strain
Alcalase (Novozymes)	*B. lichenformis*	WT	*B. lichenformis*
Savinase (Novozymes)	*B.* clausii	WT	*B. clausii*
Purafect (Genencor)	*B. lentus*	WT	*B. subtilis*
Everlase (Novozymes)	*B. clausii*	PE	*B. clausii*
Purafect OxP (Genencor)	*B. lentus*	PE	*B. subtilis*
Esperase (Novozymes)	*B. halodurans*	WT	*B. halodurans*
Kannase (Novozymes)	*B clausii*	PE	*B. clausii*
Properase (Genencor)	*B. alkalophilus*	PE	*B. alkalophilus*

[a] WT, wild type; PE, protein engineered.

most likely reflects the effects of evolution, rather than any intrinsic physical-chemical limitations on proteins (Miyazaki *et al*., 2000). Interestingly, it has also been observed that, in natural proteins, serines are statistically less prevalent in thermophilic enzymes compared to mesophilic ones (Wintrode *et al*., 2001).

Another strategy for engineering a cold-adapted subtilisin has been attempted (Tindbaek *et al*., 2004) through creating a hybrid molecule where a stable mesophilic subtilisin, savinase (Table 1), was site-directedly modified to include residues from the binding region of psychrophilic subtilisin (S39). A 12 amino acid region (MSLGSSGESSLI) of the binding cleft of S39, from Antarctic *Bacillus* TA39, was predicted to be highly flexible and was used to replace corresponding 12 residues (LSLGSPSPSATL) in savinase. The rationale being that local or global flexibility seems to be the main adaptive character of psychrophilic enzymes responsible for the thermodynamic parameters that increase the turnover at low temperature, *i.e.* decrease in activation enthalpy and increase in entropy (Lonhienne *et al*., 2000). In line with predictions, the hybrid enzyme showed the same temperature optimum and pH profile as savinase; had higher specific activity with synthetic substrates; had broader substrate specificity at ambient temperature and showed a decrease in thermostability akin to the psychrophilic enzymes.

4. APPLICATIONS OF SUBTILISIN IN DETERGENTS

The largest industrial application of enzymes is in detergents. Enzymes were first introduced into detergents early in the early 1930s. Initially the use of enzymes from animal sources led to few successes, as those enzymes were not suited to prevailing

washing conditions. A major breakthrough for detergent enzymes occurred in 1963 with the launch of alcalase (subtilisin Carlsberg from *Bacillus licheniformis* (Table 1), with a low alkaline pH optimum. Enzymes incorporated into detergents must exhibit satisfactory catalytic activities in the presence of other components and the washing conditions. Proteolytic enzymes potentially suited to use in detergents, therefore, must be stable at alkaline pH, at relatively high temperatures and in the presence of sequestering agents, bleach and surfactants. Of the various classes of proteases, only the serine proteases are potentially suited to inclusion in detergents. The bacterial subtilisins were identified, at an early stage, as being the most suitable for detergent applications. Current consumer demands together with the increased use of synthetic fibres, which do not tolerate high temperatures very well, has led to the use of lower washing temperatures. In the light of this trend coupled with the impressive bioengineering studies, *e.g.* Properase (Table 1), the applicability of subtilisins has been further enhanced.

Most industrial enzymes are produced using micro-organisms. Currently, the majority of subtilisins used in detergents are isolated from *Bacillus licheniformis, B. lentus, B. alcalophilus* or *B. amyloliquefaciens* (Subtilisin BPN'). They generally: - (1) display high activity at the pH of detergent-containing wash water; (2) are reasonably stable in the presence of other detergent components; (3) display a broad substrate specificity, rendering them capable of hydrolyzing a range of protein structures. They are produced, extracellularly, in large quantities by fermentation technology (for a pertinent review see Gupta *et al*., 2002a). They can now also be generated by recombinant (molecular biological) techniques and engineered in many respects, as already described. The literature prior to 2002 regarding various types and sources of bacterial alkaline proteases, yield improvement methods and development of novel proteases has also been reviewed (Gupta *et al*., 2002b).

Another adversary in the detergent is the presence of bleach that oxidises sensitive residues near the active sites of the subtilisins, *e.g.* methionine and cysteine. This obstacle can be overcome by site directed mutagenesis to replace the sensitive residues with ones that do not adversely affect catalytic activity, such as serine or alanine in place of methionine. This has led to the development of second-generation oxidation-resistant engineered subtilisins. Such products are Purafect OxP (Genencor) and Everlase (Novozymes), which have been on the market for some years (Table 1).

One of the first proteases used in detergents was subtilisin Carlsberg (Alcalase, Table 1) obtained from *B. licheniformis* (Fig. 1). It is a single polypeptide chain of 275 amino acids exhibiting typical Michaelis-Menten hyperbolic kinetics. Subtilisin BPN' from *B. amyloliquefaciens* was utilised at an early stage. It has 275 residues and its three-dimensional structure is very similar to that of subtilisin Carlsberg (Fig. 1), although their kinetic properties vary. Subtilisin from *B. lentus* is also used frequently as it has a better activity profile at higher pH (9–12) than subtilisin Carlsberg or BPN'. This subtilisin has 269 residues with about 60% sequence homology with each of the latter. Thus, all of the subtilisins used in detergents are of about this size, *i.e.* 27 kDa. Their significance is evinced by the fact

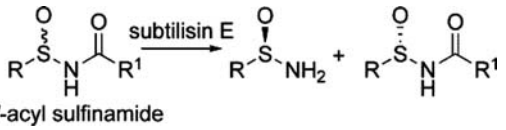
that some 900 tons of pure subtilisin were produced and used in the European Union in 2002 (Maurer, 2004). That they are produced as extracellular enzymes is a major benefit as it greatly simplifies the separation of the enzyme from the biomass and facilitates relatively straightforward downstream purification processes (Gupta *et al*., 2002a).

5. OTHER APPLICATIONS

Subtilisin BPN' is a good example of a serine protease that can also be a useful catalyst for peptide synthesis when dissolved in high concentrations of a water-miscible organic solvent such as N,N-dimethylformamide (DMF). For example, in 50% DMF, the turnover rate for peptide hydrolysis was only 1% of that in aqueous solution, whereas the turnover rate for the hydrolysis of ester substrates remained unchanged (Kidd *et al*., 1999). X-ray crystallography revealed that the imidazole ring of histidine 64 had rotated. Two new molecules of water stabilized the new conformation of the active site, with the loss of the low-barrier hydrogen bonds that had existed between histidine 64 and aspartate 32. Thus, providing a structural basis for the change in activity of these serine proteases in the presence of organic solvents.

The ability of wild type proteases, such as subtilisin, to catalyse synthetic reactions is, perhaps, surprising but not particularly efficient. Recent developments have led to very significant improvements in the applicability of subtilisin from *B. lentus* in peptide and glycopeptide syntheses (Martsumoto *et al*., 2002; Doores and Davis, 2005). A combination of site directed mutagenesis and chemical modifications with polar prosthetic groups, targeting the primary specificity pocket of the enzyme's active site, have led to very significant rate enhancement and broadening of substrate specificity. These "polar patch" or chemically modified mutants have shown remarkable utility in peptide synthesis and can also generate glycopeptides in very high yield. Another approach to improving the peptide synthetic efficiency of subtilisin is site-selective glycosylation of the active site. Again, glycosylated subtilisin from *B. lentus* had greatly increased esterase and greatly reduced amidase activities; conditions which favour formation of amide bond rather than hydrolysis (Lloyd *et al*., 2000). Glycosylation of the primary substrate binding pocket also led to a significant broadening of stereospecificity in peptide synthesis (Martsumoto *et al.*, 2001).

Recent observations extend the range of applications of subtilisin into the realm of chemical syntheses (Savile *et al*., 2005). Subtilisin E from *B. subtilis* has been identified as the most suitable hydrolase for the catalysis of the reaction shown in Scheme 1, where enantiopure arylsulfinamides ($R-S(O)-NH_2$) can be generated in gram quantities at neutral pH. These products are useful sulfinyl chiral auxilaries for synthesis of amines. These experiments highlight the stereoselectivity of enzymes since it appears that the enantioselectivity seen here arises from a favourable interaction between the aryl group of the fast-reacting (R)- arylsulfinamide and the leaving group pocket at the active site in subtilisin E.



Scheme 1. Subtilisin catalysed resolution of sulfinamides.

6. EPILOGUE

Bacterial subtilisins have served mankind well with respect to their use in detergents and they also have other proven and potential applications. The subtilisin clan has been instructive in terms of our understanding of the evolution of the structure and function of serine proteases. Considering their relatively small size (27 kDa) they have also provided molecular biologists with an excellent scaffold for protein engineering experiments. These experiments have not only generated much intellectual satisfaction but also provided us with much improved enzyme preparations through judicious directed evolution. Thus, the requirement to adjust the products to meet the needs of the modern customer has been addressed.

REFERENCES

Almog, O., Gallagher, D.T., Ladner, J.E., Strausberg, S., Alexander, P., Bryan, P. and Gilliland, G.L. (2002) Structural basis of thermostability. Analysis of stabilizing mutations in subtilisin BPN'. J Biol Chem *277*, 27553–27558.

Bryan, P.N. (2000) Protein engineering of subtilisin. Biochim Biophys Acta *1543*, 203–222.

Doores, K J. and Davis, B.G. (2005) "Polar patch" proteases as glycopeptiligases. Chem Commun 168–170.

Gros, P., Betzel, C.H., Dauter, Z., Wilson, K.S. and Hol, W.G.J. (1989) Molecular dynamics refinement of a thermitase-eglin-c complex at 1.98 Å resolution and comparison of two crystal forms that differ in calcium content. J Mol Biol *210*, 347–367.

Gupta, R., Beg, Q.K., Khan, S. and Chauhan, B. (2002a) An overview on fermentation, downstream processing and properties of microbial alkaline proteases.Appl Microbiol Biotechnol *60*, 381–395.

Gupta, R., Beg, Q.K. and Lorenz, P. (2002b) Bacterial alkaline proteases: molecular approaches and industrial applications. Appl Microbiol Biotechnol *59*, 15–32.

Jarnagin, A.S. and Ferrari, E. (1992) Extra-cellular enzymes: gene regulation and structure function relationship studies. Biotechnolog *22*, 189–217.

Kidd, R.D., Sears, P., Huang, D.H., Witte, K., Wong, C.H. and Farber, G.K. (1999) Breaking low barrier hydrogen bond in a serine protease. Protein Sci *8*, 410–417.

Lloyd, R.C., Davis, B.G. and Jones, J.B. (2000) Site-selective glycosylation of subtilisin *Bacillus lentus* causes dramatic increase in esterase activity. Bioorg Med Chem *8*, 1537–1544.

Lonhienne, T., Gerday, C. and Feller, G. (2000) Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility. Biochim Biophys Acta *1543*, 1–10.

Martsumoto, K., Davis, B.G. and Jones, J.B. (2001) Glycosylation of the primary binding pocket of a subtilisin protease causes remarkable broadening in stereospecificity in peptide synthesis. Chem Commun 903–904.

Martsumoto, K., Davis, B.G. and Jones, J.B. (2002) Chemically modified "polar patch" mutants of subtilisin in peptide synthesis with remarkably broad substrate acceptance: designing combinatorial biocatalysts. Chemistry *8*, 4129–4137.

Maurer, K-H. (2004) Detergent proteases. Curr Opin Biotechnol *15*, 330–334.

Miyazaki, K., Wintrode, P.L., Grayling, R.A., Rubingh, D.N. and Arnold, F.H. (2000) Directed evolution study of temperature adaptation in a psychrophilic enzyme. J Mol Biol *297*, 1015–1026.

Moree, W.J., Sears, P., Kawashiro, K., Witte, K. and Wong, C.H. (1997) Exploitation of subtilisin BPN' as catalyst for the synthesis of peptides containing noncoded amino acids, peptide mimetics and peptide conjugates. J Amer Chem Soc *119*, 3942–3947.

Perona, J.J. and Craik, C.S. (1995) Structural basis of substrate specificity in the serine proteases. Protein Sci *4*, 337–60.

Polgár, L. (2005) The catalytic triad of serine peptidases. CMLS, Cell Mol Life Sci *62*, 2161–2172.

Rawlings, N.D. and Barrett, A.J. (1993) Evolutionary families of peptidases. Biochem J *290*, 205–18.

Savile, C.K., Magliore, V.P. and Kazlauskas, R.J. (2005) Subtilisin-catalyzed resolution of N-Acyl Arylsulfinamides. J Am Chem Soc *127*, 2102–2113.

Shinde, U. and Inouye, M. (2000) Intramolecular chaperones; polypeptide extensions that modulate protein folding. Semin Cell Dev Biol *11*, 35–44.

Siezen, R.J. and Leunissen, J.A.M. (1997) The superfamily of subtilisin-like serine proteases. Protein Sci *6*, 501–523.

Stratton, J.R., Pelton, J.G. and Kirsch, J.F. (2001) A novel engineered subtilisin BPN' lacking a low-barrier hydrogen in the catalytic triad. Biochemistry *40*, 10411–10416.

Strausberg, S.l., Ruan, B., Fisher, K.E., Alexander, P.A. and Bryan, P.N. (2005) Directed co-evolution of stability and catalytic activity in calcium-free subtilisin. Biochemistr *44*, 3272–3279.

Subbian, E., Yabuta, Y. and Shinde, U.P. (2005) Folding pathway mediated by an intramolecular chaperone: intrinsically unstructured propeptide modulates stochastic activation of subtilisin. J Mol Biol *347*, 367–383.

Takagi, H. and Takahashi, M. (2003) A new approach for alteration of protease functions: pro-sequence engineering. Appl Microbiol Biotechnol *63*, 1–9.

Tindbaek, N., Svendsen, A., Oestergaard, P. sR. and Draborg, H. (2004) Engineering a substrate-specific cold-adapted subtilisin. Protein End Des Sel *17*, 149–156.

Wintrode, P.L., Miyazaki, K. and Arnold, F.H. (2000) Cold adaptation of a mesophilic subtilisin-like protease by laboratory evolution. J Biol Chem *275*, 31635–31640.

Wintrode, P.L., Miyazaki, K. and Arnold, F.H. (2001) Patterns of adaptation in a laboratory evolved thermophilic enzyme. Biochim Biophys Acta *1549*, 1–8.

Zhao, H. and Arnold, F.H. (1999) Directed evolution converts subtilisin E into a functional equivalent of thermitase. Protein Eng *12*, 47–53.

CHAPTER 13

ASPARTIC PROTEASES USED IN CHEESE MAKING

FÉLIX CLAVERIE-MARTÍN* AND MARÍA C. VEGA-HERNÁNDEZ

Unidad de Investigación, Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Spain

**fclamar@gobiernodecanarias.org*

1. INTRODUCTION

The use of aspartic proteases (APs) in cheese manufacture is among the earliest applications of enzymes in food processing, dating back to approximately 6000 B.C. (Fox and McSweeney, 1999). Enzymatic milk coagulation is a two-phase process. In the first phase, APs hydrolyse the Phe^{105}-Met^{106} bond of bovine κ-casein splitting the protein molecule in two, yielding hydrophobic para-κ-casein and a hydrophilic part known as the macropeptide. The second phase consists of the coagulation of the casein micelles that have been destabilized by the proteolytic attack.

Milk-clotting enzymes are obtained from mammals, plants and fungi. They can also be produced using recombinant DNA technology. Enzymes extracted from the fourth stomach (abomasum) of suckling calves (rennet) have traditionally been used as milk coagulants for cheese production. In addition to its chymosin content, conventional rennet also contains lower levels of pepsin A, the most representative peptidase of Family A1, characterized by its general proteolytic activity that makes it unsuitable for milk clotting (Harboe and Budtz, 1999). Plant and fungal milk coagulants present high levels of non-specific, heat-stable proteases the prolonged action of which cause bitterness in the cheese after a period of storage (Harboe and Budtz, 1999; Roserio *et al*, 2003). A world shortage of bovine rennet, due to the increased demand for cheese, encouraged the search for alternative milk coagulants. Research on fungal APs resulted in the production of enzymes that are inactivated at normal pasteurisation temperatures and contain low levels of non-specific proteases (Branner-Jorgensen *et al*, 1982; Yamashita *et al*, 1994; Aikawa *et al*, 2001).

In 1988, chymosin produced by recombinant DNA technology was first introduced to the dairy industry for evaluation. A few years later, scientists at Genencor

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 207–219

International were able to increase the production of chymosin in *Aspergillus niger* var. *awamori* to commercial levels (Dunn-Coleman *et al*., 1991). Presently, several recombinant chymosins such as Maxiren® produced by DSM, and Chymogen ® produced by Christian Hansen, are available on the market. Recombinant chymosin preparations are very pure and have high milk-clotting activity.

Chymosins from other mammalian species including lamb, kid goat, camel and buffalo calves are being considered as alternatives for milk clotting in the production of certain types of cheese (Mohanty *et al*., 1999; Elagamy, 2000; Rogelj *et al*., 2001; Vega-Hernández *et al*., 2004).

2. STRUCTURE OF ASPARTIC PROTEASES

Aspartic proteases (E.C. 3.4.23) are peptidases and exhibit a wide range of activities and specificities. They are present in animals, plants, fungi and viruses (Davies, 1990). APs have been linked to a variety of physiological functions including mammalian digestion of nutrients (*e.g.* chymosin, pepsin A), defence against pathogens, yeast virulence (*e.g.* candidapepsins), metastasis of breast cancer (*e.g.* cathepsin D), pollen-pistil interactions (*e.g.* cardosin A), control of blood pressure (*e.g.* renin), haemoglobin degradation by parasites (*e.g.* plasmepsins) and maturation of HIV proteins (retropepsin).

Structurally APs belong to the A1 pepsin family (Rawlings *et al*., 2004). Like other pepsin-like enzymes, APs are synthesized as preproenzymes (Fig. 1). After cleavage of the signal peptide the proenzyme is secreted and autocatalytically activated. In general the active enzymes consist of a single peptide chain of about 320-360 amino acid residues having molecular masses of 32-36 kDa. X-ray crystallographic analyses of various APs show that they are composed mostly of β-strand secondary structures arranged in a bilobal conformation (Fig. 2) (Cooper *et al*., 1990; Davies, 1990; Gilliland *et al*., 1990; Newman *et al*., 1991; Aguilar *et al*., 1997; Yang *et al*., 1997). The two lobes are homologous to each other and have evolved by gene duplication (Tang, 2004). The catalytic centre is located between the two lobes and contains a pair of aspartate residues, one in each lobe, that are essential for the catalytic activity. In most pepsin family enzymes, the catalytic Asp residues are contained in an Asp-Thr-X motif, where X is Ser or Thr. These Asp residues activate a water molecule that mediates the nucleophilic attack on the substrate peptide bond

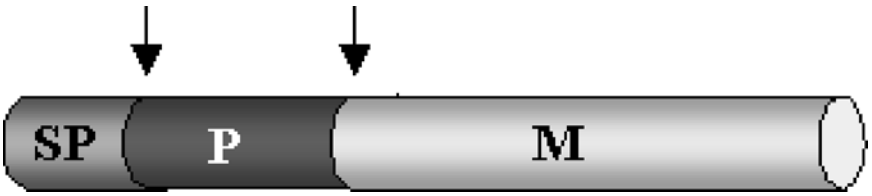


Figure 1. Schematic representation of the primary structure of bovine chymosin. SP, signal peptide; P, prosegment; M, mature enzyme. Arrows indicate processing sites

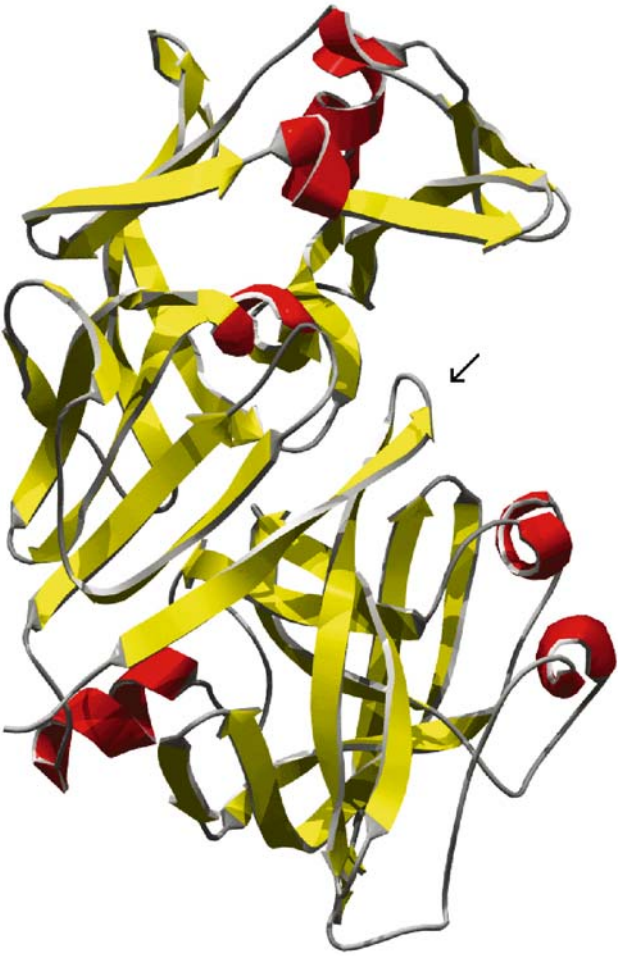


Figure 2. Tertiary structure of chymosin showing the bilobal fold. The arrow points to the flap (Gilliland *et al*, 1990)

(James, 2004). Andreeva and Rumsh (2001) have found another water molecule that plays an essential role in the formation of a chain of hydrogen bonds that determine substrate binding. The catalytic centre is large enough to accommodate at least seven residues of the polypeptidic substrate. A flexible structure (flap) located at the entrance of the catalytic site controls specificity (Hong and Tang, 2004; James, 2004). APs are active at acidic pH (Chitpinitoyl and Crabbe, 1998). It has been proposed that the optimum pH of each aspartic protease is determined by the electrostatic potential at the active site, which in turn is determined by the position and orientation of all residues near the active site (Yang *et al*, 1997).

Family A peptidases are strongly inhibited by pepstatin, a pentapeptide produced by *Streptomyces* (Marciniszyn *et al*, 1976) which contains two residues of an unusual amino acid, statine. Pepstatin binds to the flap in the catalytic site. The hydroxyl oxygen of the first statine forms hydrogen bonds with both of the catalytic aspartate residues (Davies, 1990; Yang and Quail, 1999). Pepstatin is effective against APs in general but its affinity varies between enzymes. Pepsin A is inhibited completely in the presence of equimolar amounts of pepstatin, while chymosin and gastricsin are less susceptible, 10- and 100-fold molar excesses being required for complete inhibition, respectively (Kageyama, 2002). Other inhibitors are the pepsin A inhibitor from *Ascaris*, a parasitic nematode, potato cathespin D inhibitor, and the highly selective saccaropepsin inhibitor (Kageyama, 2002).

3. CHYMOSIN

3.1. Bovine Chymosin

Chymosin (EC 3.4.23.4) is a gastric digestive aspartic peptidase that is responsible for the coagulation of milk in the abomasum of unweaned calves (Fox and McSweeney, 1999). In mammals, chymosin is expressed mostly in the foetus and the newborn, and decreases gradually during postnatal development, becoming insignificant in adults (Foltmann, 1992). The natural function of chymosin is the hydrolysis of κ-casein once the milk is in the calf's stomach, leading to the formation of a coagulum that can be easily digested. The first step in the biosynthesis of chymosin (323 amino acids, 35.6 kDa) by the cells of the gastric mucosa is the synthesis of preprochymosin, a polypeptide of 381 amino acids and 42.1 kDa (Fig. 1). Preprochymosin is secreted as an inactive precursor, known as prochymosin, with 365 amino acids and 40.8 kDa, produced by cleavage of the N-terminal signal peptide (Foltmann, 1992). In the acidic environment of the gastric lumen, prochymosin is activated by autocatalytic removal of the 42-amino acid prosegment (Pedersen *et al*, 1979). There are two allelic forms of calf chymosin, A and B, both of which are active and differ by a single amino acid substitution, Asp/Gly, at position 243 (Foltmann *et al*, 1977). The three-dimensional structure of bovine chymosin has been determined (Fig. 2) (Gilliland *et al*, 1990; Newman *et al*, 1991). Surprisingly, the native crystal structure shows that the flap at the catalytic site adopts a different conformation to that of other closely related APs such as pepsin and renin. This conformation could prevent the binding of substrate-inhibitors explaining the reduced susceptibility of chymosin to pepstatin. In addition, it could determine the specificity of chymosin to κ-casein (Gustchina *et al*, 1996). However, X-ray analysis of chymosin complexed with an inhibitor shows close resemblance to other AP-inhibitor complexes (Groves *et al*, 1998).

Bovine chymosin is used in cheese production as a milk-clotting agent because it cleaves κ-casein in a specific manner at the Phe^{105}-Met^{106} bond, and has low proteolytic activity (Fox and McSweeney, 1999; Mohanty *et al*, 1999). The yeasts *Saccharomyces cerevisiae* and *Kluyveromyces lactis,* and the filamentous fungi *Aspergillus niger* var *awamori* and *Trichoderma reesei* have been successfully used

as hosts for the expression of recombinant calf chymosin (Mohanty *et al*., 1999). Recombinant chymosin has also been produced in *Escherichia coli* but the use of this product in cheese manufacture is not accepted. Several biotechnology companies are producing the recombinant enzyme for commercial application, and different types of conventional cheeses have been produced using these preparations at experimental or pilot scales. No major differences have been detected between cheeses made with recombinant chymosin or natural enzymes regarding cheese yield, texture, smell, flavour and ripening. The absence of bovine pepsin in the recombinant preparations improves cheese yield and cheese flavour development.

3.2. Other Chymosins

Chymosin is also produced by other mammalian species such as sheep, goat, buffalo, pig, camel, humans, monkeys and rats. The characterization of chymosins from these species has been the subject of several reports (Houen *et al*., 1996; Elagamy, 2000; Rogeli *et al*., 2001; Mohanty *et al*., 2003; Vega-Hernández *et al.*, 2004). Recently, our group has cloned the cDNA for goat prochymosin and expressed it in yeast (Vega-Hernández *et al*., 2004). The cDNA encodes a protein of 381 amino acids with an N-terminal leader sequence and a proenzyme region of 16 and 42 amino acids, respectively. The deduced sequence shows high similarity to other preprochymosins (99, 94, and 94% amino acid identity with lamb, calf and buffalo sequences, respectively). The two catalytic aspartate residues of APs are conserved in the caprine sequence and the presence of six cysteine residues suggests the presence of three disulfide bridges similar to those reported for the bovine enzyme (Foltmann *et al*., 1979; Vega-Hernández *et al*., 2004). In caprine prochymosin, glutamate occupies position 36 in the propeptide. This non-conservative replacement in aspartic protease zymogens has also been observed in lamb, sheep and mouflon prochymosins (Pungercar *et al*., 1990; Francky *et al*., 2001). The recombinant caprine chymosin shows high specificity towards κ-casein and has been used experimentally to produce cheese from goat's milk (Vega-Hernández *et al.*, 2004; Vega-Hernández and Claverie-Martín, unpublished). This proteolytic capability is in agreement with the observation by Francky *et al*. (2001) that a basic residue at position 36 of prochymosin is not essential for its autocatalytic activation.

Buffalo (*Bubalos bubalis*) milk is the major milk source in India, and it has a different composition from that of cow. Mohanty *et al*. (2003) have purified chymosin, (molecular weight of 35.6 kDa) from the stomach of buffalo calves. Slight differences in stability and relative proteolytic activity are found compared to bovine chymosin. This indicates that buffalo chymosin could be the best choice for cheese production from buffalo milk.

Studies on the characteristics of rennet extracted from camel stomach (*Camelus dromedaries*) have been reported (Elagamy, 2000). Camel chymosin has a specific κ-casein hydrolysis activity superior to that of bovine chymosin (Kappeler *et al*., 2004). Consequently, in cheese made with camel chymosin the loss of protein due to non-specific degradation is decreased, yield is improved and

the development of bitter taste is reduced. Were camel chymosin to be commercially available, more efficient clotting of camel's milk could be achieved at the industrial level. Furthermore, camel chymosin is very suitable for the coagulation of bovine milk.

Lamb preprochymosin cDNA has been cloned and expressed in *E. coli* and the recombinant lamb chymosin has been tested for its potential use in cheese production (Rogelj *et al*, 2001). The coagulation properties of recombinant lamb chymosin and the overall quality of the cheese made with this enzyme are similar to those of recombinant bovine chymosin. A characteristic of recombinant lamb chymosin is its instability at temperatures above 45 °C (Rogelj *et al*, 2001). This could be an advantage in the production of hard cheeses where relatively high incubation temperatures are used. The production of cheeses made from ovine milk is also a potential area for the application of recombinant lamb chymosin.

4. PLANT ASPARTIC PROTEASES

Similarly to other aspartic proteases, plant APs are synthesized as single-chain zymogens. Subsequent maturation is a crucial step in the regulation of their activity. The primary structure of plant APs comprises a signal peptide, responsible for translocation to the endoplasmic reticulum; a prosegment of 46-50 amino acids involved in correct folding, stability and sorting of the enzyme (Simöes and Faro, 2004); and the mature enzyme which possesses two catalytic sequence motifs (Fig. 3). In contrast to other APs, the two catalytic Asp residues are contained within Asp-Thr-Gly and Asp-Ser-Gly motifs (Simöes and Faro, 2004). Plant APs contain an extra region of approximately 100 amino acids named the plant-specific insert (PSI) that presents no homology to any other aspartic protease sequence. The PSI is usually removed during the maturation process and resembles saposin-like proteins (SALIPS). The function of PSI is still unclear but a possible role in vacuolar targeting has been proposed (Egas *et al*, 2000).

Most plant APs are located in seeds (suggesting a role in storage-protein cleavage), in leaves (indicating a role in mechanisms of defence against pathogens),

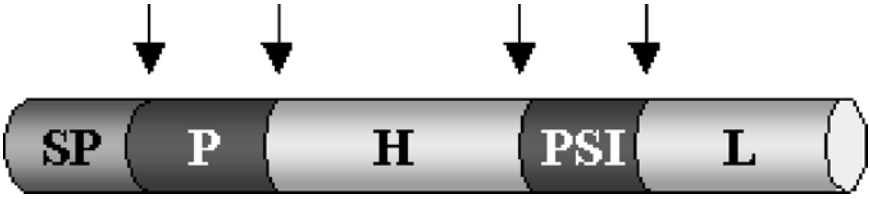


Figure 3. Schematic representation of the primary structure of cardosin A. SP, signal peptide; P, prosegment; H, heavy chain of the mature enzyme; L, light chain of mature enzyme; PSI, plant-specific insert. Arrows indicate processing sites

or in flowers (implying a role in sexual reproduction) (Simões and Faro, 2004). Plant APs are also involved in defence mechanisms and cell death events associated with plant senescence and response to stress.

Plant extracts have been used as coagulants in cheese making for many centuries (Roserio *et al*., 2003). In contrast to chymosin which is specific for κ-casein, the APs present in plant extracts cleave α-, β- and κ-caseins. This causes excessive acidity, bitterness and texture defects in cheese, thereby limiting their use. However, these characteristics are responsible for the special flavour, smell and consistency of the cheese varieties produced using plant enzymes. Cheeses made with plant coagulants are found mainly in Southern European and West African countries. Flower extracts of cardoon and red star thistle (*Cynara* sp. and *Centaurea calcitrapa*) are used in Portugal and Spain for the manufacture of traditional cheeses (Roserio *et al*., 2003). The main milk-clotting APs present in these extracts are known as cardosins, cyprosins and cenprosins (Ramalho-Santos *et al*., 1997; White *et al*., 1999; Domingos *et al*., 2000). There are other APs isolated from *Cynara* sp., referred to as cynarases, but they have not been well characterized (Roserio *et al*., 2003; Sidrach *et al*., 2004). In Nigeria, extracts from the Sodom apple (*Calotropis procera*) are used in the production of traditional cheese. Plant recombinant APs, including cardosins and cyprosins, have been expressed in yeast but these enzymes are not yet commercially available for industrial application (Soares Pais *et al*., 2000; Castanheira *et al*., 2005).

4.1. Cardosins

Cardosin A, the most abundant of the cardosins, accumulates in the protein storage vacuoles of the stigmatic epidermal papillae and in the vacuoles of the epidermal cells in the stylus (Ramalho-Santos *et al*., 1997). Preprocardosin A is encoded by the *CARDA* gene and consists of 504 amino acid residues (Fig. 3). The mature enzyme is formed by two peptides of 31 and 15 kDa and has low proteolytic activity (Roserio *et al*., 2003). The conversion to the active enzyme probably takes place inside the vacuoles, during which process the PSI is removed prior to cleavage of the prosegment (Ramalho-Santos *et al*., 1998). The catalytic residues are located in positions 32 and 215 of the heavy chain (Frazao *et al*., 1999). The unique feature of cardosin A among plant APs is the presence of the RGD cell attachment motif Arg^{176}-Gly^{177}-Asp^{178} (Frazao *et al*., 1999) which is a well-known integrin-binding sequence. It has been suggested that this enzyme may participate in an RGD-dependent proteolytic mechanism in pollen-pistil interactions (Simões and Faro, 2004). The C-terminus of cardosin A contains the hydrophobic sequence Val^{320}-Gly^{321}-Phe^{322}-Ala^{323}-Glu^{324}-Ala^{325}-Ala^{326} which is conserved among plant APs and has been proposed to play a role in vacuolar targeting (Frazao *et al*., 1999; Ramalho-Santos *et al*., 1998). The crystal structure of cardosin A has been determined and shows the bilobal structure observed for other APs (Frazao *et al*., 1999). The two independent polypeptides are held together by hydrophobic interactions and hydrogen bonds.

In contrast to cardosin A, cardosin B accumulates in the cell wall and in the extracellular matrix of the transmitting tissue, suggesting that the two cardosins may play different roles in the pistil of *Cynara cardunculus* (Vieira *et al*., 2001). Although less abundant than Cardosin A, cardosin B has more proteolytic activity and may take part in general protein digestion (Faro *et al*., 1995). This enzyme displays 73% similarity to cardosin A (Vieira *et al*., 2001). The cardosin B precursor (506 amino acids) is encoded by the *CARDB* gene. The mature enzyme is formed by two peptides of 34 kDa and 14 kDa, the catalytic residues being located in the former (Vieira *et al*., 2001). Cardosin B lacks the RGD motif and an additional putative N-glycosylation site is created by the replacement of an Asp by Asn.

4.2. Cyprosins

Cyprosins are another type of AP isolated from the flowers of *C. cardunculus* (Ramalho-Santos *et al*., 1997). Preprocyprosin (509 amino acid residues) is encoded by the *CYPRO1* gene. The precursor of recombinant cyprosin is processed to different isoforms by excision of the prosegment and most of the PSI (White *et al.*, 1999). The mature enzyme shares 52% identity with animal cathepsin D, the closest AP of non-plant origin (White *et al*., 1999).

5. FUNGAL ASPARTIC PROTEASES

Due to the worldwide shortage of calf chymosin, fungal APs have been being used as milk-clotting enzymes in the dairy industry for about 30 years. The enzymes produced by *Mucor miehei, Mucor pusillus* and *Cryphonectria (Endothia) parasitica*, marketed under the trade names Rennilase®, Fromase®, Novoren®, Marzyme®, Hannilase®, Marzyme® and Suparen®, are widely used for the production of different kinds of cheese. The proteolytic specificities of the fungal coagulants are different from those of calf chymosin, resulting in the production of some bitter peptides during the process of cheese ripening (Harboe and Budtz, 1999).

The crystal structures of several fungal APs including those of *R. miehei*, *R. pusillus*, *C. parasitica* and *Irpex lacteus*, alone and complexed with inhibitors, have been determined (Blundell *et al*., 1990; Newman *et al*., 1993; Yang *et al*., 1997; Yang and Quail, 1999; Coates *et al*., 2002; Fujimoto *et al*., 2004). The tertiary structures of these enzymes are very similar to those of other APs.

5.1. Mucorpepsin

The APs produced extracellularly by the two closely related species of zygomycetes, *M. pusillus* and *M. miehei*, possess relatively high milk-clotting activities due to their selective cleavage of κ-casein, along with relatively low proteolytic activities (Harboe and Budtz, 1999; Awad *et al*., 1999). These enzymes, referred to as mucorpepsins or rennins, exhibit the highest levels of thermal stability among

the APs and therefore are of limited use as milk coagulants. This high thermal stability results in the persistence of enzyme activity after cooking of the curd and thus causing off-flavours in the cheese during long maturation periods (Yamashita *et al.*, 1994).

The aspartic protease from *M. miehei* (MAP, EC 3.4.23.23) is the most glycosylated of the AP enzymes (Rickert and McBride-Warren, 1974). The carbohydrate moieties may stabilize the conformation of MAP, conferring the enzyme a high level of thermal stability and protecting it from proteolytic attack (Yang *et al.*, 1997). MAP was first crystallized by Jia *et al.* (1995), and Yang *et al.* (1997) refined the native enzyme structure. Based on this structure, useful modifications of the enzyme resulting in reduced thermal stability and higher milk-clotting activity have been designed.

MAP preparations used in the cheese industry have been produced both in *M. miehei* and in *A. oryzae* (Budtz and Heldt-Hansen, 1998; Harboe and Kristensen, 2000). When these preparations are deglycosylated by treatment with endo-β-N-acetylglucosaminidase H a significant increase in clotting activity is observed (Harboe and Kristensen, 2000). Two procedures have been described to reduce the thermal stability of MAP; one by treatment in aqueous solution with oxidizing agents containing active chlorine, and the other by acylation with an active derivative of a carboxylic acid (Branner-Jorgensen, 1981; Branner-Jorgensen *et al.*, 1982).

As expected, *M. miehei* MAP is almost identical to that of *M. pusillus* AP (MPAP) (Tonouchi *et al.*, 1986; Yang *et al.*, 1997). While they share a common antigenic structure and almost identical enzymatic properties, there are some differences between these two enzymes with respect to their peptide cleavage patterns and glycosylation. The structural gene for the *M. pusillus* aspartic protease, *mpr*, has been expressed in *S. cerevisiae* (Aikawa *et al.*, 1990). The product secreted to the culture medium is the proenzyme: The 44 amino acid prosegment is removed by autocatalytic processing at acid pH (Hiramatsu *et al.*, 1989).

Deglycosylation studies have shown that removal of the N-linked carbohydrate groups from MPAP increases its milk-clotting activity whilst decreasing both proteolytic activity and thermal stability (Aikawa *et al.*, 1990). Site-directed mutagenesis at several positions has been carried out in order to improve its practical properties in cheese production. Yamashita *et al.* (1994) have obtained mutant forms of this aspartic protease having decreased thermal stability. Mutant *mpr* genes carrying Gly186Asp or Ala101Thr have been expressed in *S. cerevisiae*. Both mutations cause a marked decrease in thermal stability of the enzyme. The double mutant shows the lowest thermal stability without affecting the enzymatic activity (Yamashita *et al.*, 1994). By contrast, replacement of Tyr^{75} in the flap by Asn has been shown to reduce the non-specific proteolytic activity of this enzyme, leading to a considerable enhancement of the specific clotting activity (Park *et al.*, 1996). In addition, mutant Glu13Ala shows a 5-fold increase in the ratio of clotting versus proteolytic activity without significant loss of clotting activity (Aikawa *et al.*, 2001). Residue Glu^{13} seems to play a critical role in forming the correct hydrogen bond network around the active centre.

5.2. Endothiapepsin

The chestnut blight fungus *C. parasitica* produces an aspartic protease, endothiapepsin, with milk-clotting properties similar to those of calf rennet (Awad *et al.*, 1999). Due to its very high thermolability this enzyme is particularly suited for use in the production of Emmental and Italian style cheeses. Extensive production of a wide variety of cheeses including Cheddar, Swiss, Colby and Italian varieties manufactured with partially purified preparations of *C. parasitica* enzyme are considered to be equal or superior to control cheeses made with animal rennet. DSM markets Endothiapepsin under the trade name Suparen®.

C. parasitica protease has greater proteolytic activity than MAP or chymosin. Proteolysis of both α_{s1}-casein and β-casein occurs during storage of Mozzarella and Cheddar cheeses made with *C. parasitica* protease (Yun *et al*, 1993). The change in cheese properties during storage is related to the combined effects of hydrolysis of α_{s1}-casein and β-casein. Kim *et al*. (2004) have combined the use of chymosin and the *C. parasitica* enzyme to control the hydrolysis of α_{s1}-casein and β-casein during aging of Cheddar cheese to independently control its firmness and meltability while regulating undesirable levels of bitterness associated with high levels of *C. parasitica* protease.

5.3. Irpex Lacteus Protease

The wood-decaying basidiomycete *Irpex lacteus* produces an AP (ILAP) that has high milk-clotting activity in relation to its proteolytic activity. This enzyme may become a good chymosin alternative (Kobayashi *et al*, 1985). ILAP contains 340 amino acid residues with a molecular mass of 35 kDa. It is most active at pH 3.0 and is inhibited by pepstatin. A feature of ILAP is its high content of serine and threonine residues (48 and 54, respectively), accounting for 30% of the residues which is double the average for other proteins. It lacks the three-disulfide bridges that are generally present in most pepsin-type APs.

REFERENCES

Aguilar, C.F., Cronin, N.B., Badasso, M., Dreyer, T., Newman, M.P., Cooper, J.B., Hoover, D.J., Wood, S.P., Johnson, M.,S. and Blundell, T. (1997) The three-dimensional structure at 2.4Å resolution of glycosylated proteinase A from the lysosome-like vacuole of *Saccharomyces cerevisiae* J. Mol. Biol. *267*, 899–915.

Aikawa, J., Park, Y.N., Sugiyama, M., Nishiyama, M., Horinouchi, S. and Beppu, T. (2001) Replacement of amino acid residues at subsites and their effects on the catalytic properties of *Rhizomucor pusillus* pepsin, an aspartic proteinase from *Rhizomucor pusillus*. J. Biochem. *129*, 791–794.

Aikawa, J., Yamashita, T., Nishiyama, M., Horinouchi, S. and Beppu, T. (1990) Effects of glycosylation on the secretion and enzyme activity of Mucor rennin, an aspartic proteinase of *Mucor pusillus*, produced by recombinant yeast. J. Biol. Chem. *265*, 13955–13959.

Andreeva, N. and Rumsh, L.D. (2001) Analysis of crystal structures of aspartic proteinases: On the role of amino acid residues adjacent to the catalytic site of pepsin-like enzymes. Protein Sci. *10*, 2439–2450.

Awad, S., Luthi-Peng, Q.Q. and Puhan, Z. (1999) Proteolytic activities of Suparen and Rennilase on buffalo, cow, and goat whole casein and beta-casein. J. Agric. Food Chem. *47*, 3632–3639.

Blundell, T.L., Jenkins J.A., Sewell, B.T., Pearl, L.H., Cooper, J. B., Tickle, I.J., Veerapandian, B. and Wood, S.P. (1990) X-ray analyses of aspartic proteinases. The three-dimensional structure at 2.1 Å resolution of endothiapepsin. J. Mol. Bio941.

Branner-Jorgensen, S., Inventor; Branner-Jorgensen, S., assignee. (1981 March 10) Thermal destabilization of microbial rennet. U.S. Patent 4,255,454.

Branner-Jorgensen, S., Schneider, P. and Eigtved, P., inventors; Novo Industri A/S, assignee. (1982 November 2) Thermal destabilization of microbial rennet. U.S. Patent 4, 357, 357.

Budtz, P. and Heldt-Hansen, H. P., inventors; Gist-brocades, B. V., assignee. (1998 September 1) Cheese making with recombinant aspartic protease. U.S. Patent 5, 800, 849.

Castanheira, P., Samyn, B., Sergeant, K., Clemente, J.C., Dunn, B. M., Pires, E. Van Beeumen, J. and Faro., C. (2005) Activation, proteolytic processing, and peptide specificity of recombinant cardosin A. J. Biol. Chem. *280*, 13047–13054

Chitpinitoyl, S. and Crabbe, M.J.C. (1998) Chymosin and aspartic proteinases. Food Chem. *61*, 395–418.

Coates, L., Erskine, P.T., Crump, M.P., Wood, S.P. and Cooper, J.B. (2002) Five atomic resolution structures of endothiapepsin inhibitor complexes: Implications for the aspartic proteinase mechanism. J. Mol. Biol. *318*, 1405–415.

Cooper, J.B., Khan, G., Taylor, G., Tickle, I.J. and Blundell, T. (1990) X-ray analysis of aspartic proteinases. II. Three-dimensional structure of the hexagonal crystal form of porcine pepsin at 2.3 Å resolution. J. Mol. Biol. *214*, 199–222.

Davies, D.R. (1990) The structure and function of the aspartic proteinases. Annu. Rev. Biophys. Biophys. Chem. *19*, 189–215.

Domingos, A., Cardoso, P.C., Xue, Z.T., Clemente, A., Brodelius, P.E. and Pais, M.S. (2000) Purification, cloning and autoproteolytic processing of an aspartic proteinase from *Centaurea calcitrapa*. Eur. J. Biochem. *267*, 6824–6831.

Dunn-Coleman, N.S., Bloebaum, P., Berka, R.M., Bodie, E., Robinson, N., Armstrong, G., Ward, M., Przetak, M., Carter, G.L., LaCost, R., Wilson, L.J., Kodama, H.K., Baliu, E.F., Bower, B., Lamsa, M. and Heinsohn, H. (1991) Commercial levels of chymosin production by *Aspergillus*. Biotechnol. *9*, 976–891.

Elagamy, E.I. (2000) Physicochemical, molecular and immunological characterization of camel calf rennet: A comparison with buffalo rennet. J. Dairy Res. *67*, 73–81.

Egas, C., Lavoura, N., Resende, R., Brito, R.M.M., Pires, E., Pedroso de Lima, M.C. and Faro, C. (2000) The saposin-like domain of the plant aspartic proteinase precursor is a potent inducer of vesicle leakage. J. Biol. Chem. *49*, 38190–38196.

Faro, C., Verissimo, P., Lin, Y., Tang, J. and Pires, E. (1995) Cardosin A and B, aspartic proteases from the flowers of cardoon. Adv. Exp. Med. Biol. *362*, 373–377.

Foltmann, B. (1992) Chymosin: a short review on foetal and neonatal gastric proteases. Scand. J. Clin. Lab. Invest. Suppl. *210*, 65–79.

Foltmann, B., Barkholt, P., Kauffman, D. and Wybrandt, G. (1979) The primary structure of calf chymosin. J. Biol. Chem. *254*, 8447–8451.

Foltmann, B., Pedersen, V.B., Jacobsen, H., Kauffman, D. and Wybrandt, G. (1977) The complete amino acid sequence of prochymosin. Proc. Natl. Acad. Sci. U S A. *74*, 2321–2324.

Fox, P.F. and McSweeney, P.L.H. (1999) Rennets: their role in milk coagulation and cheese ripening. In Microbiology and Biochemistry of Cheese and Fermented Milk. Law, B.A. ed. Blackie Academic and Proffesional, London, pp.1–49.

Francky, A., Francky, B.M., Strukelj, B., Gruden, K., Ritonja, A., Krizaj, I., Kregar, I., Pain, R. H. and Pungercar, J. (2001) A basic residue at position 36p of the propeptide is not essential for the correct folding and subsequent autocatalytic activation of prochymosin. Eur. J. Biochem. *268*, 2362–2368.

Frazao, C. Bento, I., Costa J., Soares, C.M., Veríssimo, P., Faro, C., Pires, E., Cooper, J. and Larrondo, M.A. (1999) Crystal structure of cardosin A, a glycosilated and Arg-Gly-Asp-containing aspartic proteinase from the flowers of *Cynara cardunculus* L. J. Biol. Chem. *274*, 27694–27701.

Fujimoto, Z., Fujii, Y., Kaneko, S., Kobayashi, H. and Mizuno, H. (2004) Crystal structure of aspartic proteinase from *Irpex lacteus* in complex with inhibitor pepstatin. J. Mol. Biol. *341*, 1227–1235.

Gilliland, G.L., Winborne, E.L., Nachman, J. and Wlodawer, A. (1990) The three-dimensional structure of recombinant bovine chymosin at 2.3 Å resolution. Proteins. *8*, 82–101.

Groves, M.R., Dhanaraj, V., Badasso, M., Nugent, P., Pitts, J.E., Hoover, D.J. and Blundell, T.L. (1998) A 2.3 Å resolution structure of chymosin complexed with a reduced bond inhibitor shows that the active site beta-hairpin flap is rearranged when compared with the native crystal structure. Protein Eng. *11*, 833–840.

Gustchina, E., Rumsh, L., Ginodman, L., Majer, P. and Andreeva, N. (1996) Post X-ray crystallographic studies of chymosin: the existence of two structural forms and the regulation of activity by the interaction with the histidine-proline cluster of kappa-casein. FEBS Lett. *379*, 60–62.

Harboe, M.K. and Budtz, P. (1999) The production, action and application of rennet and coagulants. In Tecnology of cheese making. Law, B. A. ed. Sheffield Academic Press, Sheffield,pp. 33–65.

Harboe, M.K. and Kristensen, P.B., inventors; Chr. Hansen A/S, assignee. (2000 October 3) Microbially derived enzymes having enhanced milk clotting activity and method producing the same. US Patent 6, 127, 142.

Hiramatsu, R., Aikawa, J., Horinouchi, S. and Beppu, T. (1989) Secretion by yeast of the zymogen form of Mucor rennin, an aspartic proteinase of *Mucor pusillus*, and its conversion to the mature form. J. Biol. Chem. *264*, 16862–16866.

Houen, G., Madsen, M.T., Harlow, K.W., Lonblad, P. and Foltmann, B. (1996) The primary structure and enzymic properties of porcine prochymosin and chymosin. Int. J. Biochem. Cell Biol. *28*, 667–675.

Hong, L. and Tang, J. (2004) Flap position of free memapsin 2 (beta-secretase), a model for flap opening in aspartic protease catalysis. Biochem. *43*, 4689–4695.

James, M.N.G. (2004) Catalytic pathway of aspartic peptidases. In Handbook of Proteolytic Enzymes. Barrett, A.J., Rawlings, N.D. and Woessner, J.F. eds. Elsevier, London, pp. 12–19.

Jia, Z., Vandonselaar, M., Schneider, P. and Quail, J.W. (1995) Crystallization and preliminary X-ray structure solution of *Rhizomucor miehei* aspartic proteinase. Acta Crystallogr. D. *51*, 243–244.

Kageyama, T. (2002) Pepsinogens, progastricsins, and prochymosins: structure, function, evolution, and development. Cell Mol. Life Sci. 59, 288–306.

Kappeler, S., Farah, Z., van den Brink, J.M., Rahbeck-Nielsen, H., Budtz, P., inventors; (2004 September 16) Method of producing non-bovine chymosin and use hereof. United States Patent Application 20040180410.

Kim, S.Y., Gunasekaran, S. and Olson, N.F. (2004) Combined use of chymosin and protease from *Cryphonectria parasitica* for control of meltability and firmness of cheddar cheese. J. Dairy Sci. *87*, 274–283.

Kobayashi, H., Kusakabe, I. and Murakami, K. (1985) Milk-clotting enzyme from *Irpex lacteus* as a calf rennet substitute for Cheddar cheese manufacture. Agric. Biol. Chem. *49*, 1605–1609.

Marciniszyn, J. Jr., Hartsuck, J.A. and Tang, J. (1976) Mode of inhibition of acid proteases by pepstatin. J. Biol. Chem. *251*, 7088–7094.

Mohanty, A.K., Mukhopadhyay, U.K., Grover, S. and Batish, V.K. (1999) Bovine chymosin: Production by rDNA technology and application in cheese manufacture. Biotechnol. Adv. *17*, 205–217.

Mohanty, A.K., Mukhopadhyay, U.K., Kaushik, J.K., Grover, S. and Batish, V.K. (2003) Isolation, purification and characterization of chymosin from riverine buffalo (*Bubalos bubalis*). J. Dairy Res. *70*, 37–43.

Newman, M., Safro, M., Frazao, C., Khan, G., Zdanov, A., Tickle, I.J., Blundell, T.L. and Andreeva, sN. (1991) X-ray analyses of aspartic proteinases. IV. Structure and refinement at 2.2 Å resolution of bovine chymosin. J. Mol. Biol. *221*, 1295–1309.

Newman, M., Watson, F., Roychowdhury, P., Jones, H., Badasso, M., Cleasby, A., Wood, S.P., Tickle, I.J. and Blundell, T.L. (1993) X-ray analyses of aspartic proteinases. V. Structure and refinement at 2.0 A resolution of the aspartic proteinase from *Mucor pusillus*. J. Mol. Biol. *230*, 260–283.

Park, Y.N., Aikawa, J., Nishiyama, M., Horinouchi, S. and Beppu, T. (1996) Involvement of a residue at position 75 in the catalytic mechanism of a fungal aspartic proteinase, *Rhizomucor pusillus* pepsin. Replacement of tyrosine 75 on the flap by asparagines enhances catalytic efficiency. Protein Eng. *9*, 869–875.

Pedersen, V.B., Christensen, K.A. and Foltmann, B. (1979) Investigations on the activation of bovine prochymosin. Eur. J. Biochem. *94*, 573–80.

Pungercar, J., Strukelj, B., Gubensek, F., Turk, V. and Kregar, I. (1990) Complete primary structure of lamb preprochymosin deduced from cDNA. Nucleic Acids Res. *18*, 4602.

Ramalho-Santos, M., Pissarra, J.,Veríssimo, P., Pereira, S., Salema, R., Pires, E. and Faro, C.J. (1997) Cardosin A, an abundant aspartic proteinase, accumulates in protein storage vacuoles in the stigmatic papillae of *Cynara cardunculus* L. Planta. *203*, 204–212.

Ramalho-Santos, M., Veríssimo, P., Cortes, L., Samyn, B. Van Beeumen, J., Pires, E. and Faro, C. (1998) Identification and proteolytic processing of procardosin A. Eur. J. Biochem. *255*, 133–138.

Rawlings, N.D., Tolle, D.P. and Barret, A.J. (2004) MEROPS: The peptidase database. Nucleic Acid Res. 32 Database issue, D160-D164. http://www.merops.sanger.ac.uk/

Rickert, W.S. and McBride-Warren, P.A. (1974) Structural and functional determinants of *Mucor miehei* protease. IV. Nitration and spectrophotometric titration of tyrosine residues. Biochim. Biophys. Acta. *371*, 368–378.

Rogelj, I., Perko, B., Francky, A., Penca, V. and Pungercar, J. (2001) Recombinant lamb chymosin as an alternative coagulating enzyme in cheese production. J. Dairy Sci. *84*, 1020–1026.

Roserio, L.B., Barbosa, M., Ames, J.M. and Wilbey R.A. (2003) Cheesemaking with vegetable coagulants-the use of *Cynara* L. for the production of ovine milk cheeses. Int. J. Dairy Tech. *56*, 76–85.

Sidrach, L., García-Cánovas, F., Tudela, J. and Rodríguez-López, J. N. (2004) Purification of cynarases from artichoke (*Cynara scolymus* L.): enzymatic properties of cynarase A. Phytochem. *66*, 41–49.

Simöes, I. and Faro, C. (2004) Structure and function of plant aspartic proteinases. Eur. J. Biochem. *271*, 2067–2075.

Soares Pais, M.S., Calixto, F.C. and Planta, R.J. inventors; Instituto de Ciencia Aplicada e Technologia, assignee. (2000 December 14) Production by yeast of aspartic proteinases from plant origin. International Patent WO 00/75283 A1.

Tang, J. (2004) Pepsin A. In Handbook of Proteolytic Enzymes. Barrett, A.J., Rawlings, N.D. and Woessner, J.F. eds. Elsevier, London, pp. 19–28.

Tonouchi, N., Shoun, H., Uozumi, T. and Beppu, T. (1986) Cloning and sequencing of a gene for Mucor rennin, an aspartate protease from *Mucor pusillus*. Nucleic Acids Res. *14*, 7557–7568.

Vega-Hernández, M.C., Gómez-Coello, A., Villar, J. and Claverie-Martín, F. (2004) Molecular cloning and expression in yeast of caprine prochymosin. J. Biotechnol. *114*, 69–79

Vieira, M., Pissarr, J., Veríssimo, P., Castanheira, P., Costa, Y., Pires, E. and Faro, C. (2001) Molecular cloning and characterization of cDNA encoding cardosin B, an aspartic proteinase accumulating extracellularly in the transmitting tissue of *Cynara cardunculus* L. Plant Mol. Biol. *45*, 529–539.

White, P.C., Cordeiro, M.C., Arnold, D., Brodelius, P. and Kay, J. (1999) Processing, activity, and inhibition of recombinant cyprosin, an aspartic proteinase from cardoon (*Cynara cardunculus*). J. Biol. Chem. *274*, 16685–16693.

Yamashita, T., Higashi, S., Higashi, T., Machida, H., Iwasaki, S., Nishiyama, M. and Beppu, T. (1994) Mutation of a fungal aspartic proteinase, *Mucor pusillus* rennin, to decrease thermostability for use as a milk coagulant. J Biotechnol. *32*, 17–28.

Yang, J. and Quail, J.W. (1999) Structure of the *Rhizomucor miehei* aspartic proteinase complexed with the inhibitor pepstatin A at 2.7 Å resolution. Acta Crystallogr. D. Biol. Crystallogr. *55*, 625–630.

Yang, J., Teplyakov, A. and Quail, J.W. (1997) Crystal structure of the aspartic proteinase from *Rhizomucor miehei* at 2.15- Å resolution. J. Mol. Biol. *268*, 449–459.

Yun, J.J., Kiely, L.J., Kindstedt, P.S. and Barbano, D.M. (1993) Mozzarella cheese: impact of coagulant type on functional properties. J. Dairy Sci. *76*, 3657–3663.

CHAPTER 14

METALLOPROTEASES

JOHANNA MANSFELD*

Department of Biochemistry, Institute of Biotechnology, Martin-Luther University Halle-Wittenberg, Halle, Germany

**mansfeld@biochemtech.uni-halle.de*

1. INTRODUCTION: CLASSIFICATION OF METALLOPROTEASES

Metalloproteases (metallopeptidases or metalloproteinases) represent an extensive class of hydrolases which cleave peptide bonds by the action of a water molecule which is activated by complexing to bivalent metal ions. Most metalloproteases are characterised by a catalytic zinc ion. However, in some enzymes this function is undertaken by manganese, cobalt, nickel or even copper ions. In some metalloproteases two metal ions act co-catalytically. The metal ion is complexed by three conserved amino acid residues that can be His, Asp, Glu or Lys.

According to the classification of proteases based on protein structure and homology implemented in the MEROPS database (http://merops.sanger.ac.uk) (see Rawlings *et al*., 2004; Barrett *et al*., 2004; Rawlings *et al*., Chapter 10, this volume), metalloproteases are found in 14 different clans. In addition, clan M- contains metalloprotease families not yet assigned to a clan. Proteases from clans MA, MC, MD, ME, MJ, MK, MM, MO and MP require only one catalytic metal ion, in most cases zinc ions, whereas clans MF, MG, MH, MN and MQ contain two metal ions acting co-catalytically on the substrate (M stands for metalloprotease).

Clan MA, comprising subclans MA(E) (gluzincins) and MA(M) (metzincins), is one of the most comprehensive clans and contains some of the most prominent and industrially relevant members of metalloproteases which will be discussed in detail in the following sections. All members of this clan are characterised by a zinc ion at the active site and the highly conserved HEXXH motif which is integrated into the central α-helix. This motif is also found in proteins other than metalloproteases. The two His residues of this motif are involved in zinc binding, whereas Glu is considered to be the catalytically important amino acid

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 221–242

residue. Subclans MA(E) and MA(M) differ in the third zinc liganding amino acid residue. In the case of the gluzincins this ligand is a Glu residue that is 18–72 amino acid residues distant from the HEXXH motif. The following families belong to the gluzincins (a representative member of the family and its origin are given in parentheses): M1 (aminopeptidase N, *Homo sapiens*), M2 (angiotensin-converting enzyme peptidase unit 1, *Homo sapiens*), M3 (thimet oligopeptidase, *Rattus norvegicus*), M4 (thermolysin, *Bacillus thermoproteolyticus*), M5 (mycolysin, *Streptomyces cacaoi*), M9 (microbial collagenase, *Vibrio alginolyticus*), M13 (neprilysin, *Homo sapiens*), M26 (IgA1-specific metalloendopeptidase, *Streptococcus sanguis*), M27 (tentoxilysin, *Clostridium tetani*), M30 (hyicolysin, *Staphylococcus hyicus*), M32 (carboxypeptidase Taq, *Thermus aquaticus*), M34 (anthrax lethal factor, *Bacillus anthracis*), M36 (fungalysin, *Aspergillus fumigatus*), M41 (FtsH peptidase, *Escherichia coli*), M48 (Ste24 peptidase, *Saccharomyces cerevisiae*), M56 (BlaR1 peptidase, *Staphylococcus aureus*), M60 (enhancin, *Lymantria dispar* nucleopolyhedrovirus), M61 (glycyl aminopeptidase, *Sphingomonas capsulata*) (Barrett *et al*., 2004). The metzincins contain either an Asp or a His residue in an extended HEXXHXXGXXH/D motif as the third zinc ligand and in addition are characterised by a conserved Met-turn (Stöcker *et al*., 1995) underlying the active site. The conserved glycine in this motif is important for the formation of the β-turn that brings the three zinc ligands into the required position (Bode *et al.*, 1992). The catalytic mechanism of the metzincins is not as well known as that of the gluzincins. Some metzincins such as meprin, astacin and serralysin are able to hydrolyse peptide nitroanilides. The following families belong to the MA(M) metzincin subclan: M6 (immune inhibitor A, *Bacillus thuringiensis*), M7 (snapalysin, *Streptomyces lividans*), M8 (leishmanolysin, *Leishmania major*), M10 (matrix metallopeptidase-1, *Homo sapiens*), M11 (gametolysin, *Chlamydomonas reinhardtii*), M12 (astacin, *Astacus astacus*), M35 (deuterolysin, *Aspergillus flavus*), M43 (cytophagalysin, *Cytophaga* sp.), M57 (prtB gene product, *Myxococcus xanthus*), M64 (IgA protease, *Clostridium ramosum*), M66 (StcE peptidase *Escherichia coli*), and M72 (peptidyl-Asp metalloendopeptidase, *Pseudomonas aeruginosa*).

Clan MC (family M14) contains a number of carboxypeptidases, *e.g.* the important animal enzymes carboxypeptidases A, B, D and E, as well as carboxypeptidase T from actinomycetes having the conserved motif HXXE.

Clan MD contains families M15 and M74. Family M15 includes the zinc D-Ala-D-Ala carboxypeptidase from *Streptomyces albus* which releases the D-amino acid-containing cross-linking peptide (required for bacterial cell wall biosynthesis) from its precursor, the VanX D-Ala-D-Ala dipeptidase and the VanY D-Ala-D-Ala carboxypeptidase from *Enterococcus* (involved in vancomycin resistance), and endolysins from bacteriophages A118 and A500. Family M74 (containing the murein endopeptidase (MepA) from *Escherichia coli*) hydrolyses the murein crosslinks in bacterial cell walls.

Members of clan ME (formed by families M16 which contains the mitochondrial-processing peptidase MPP and M44, containing the vaccinia virus polyprotein

processing endopeptidase, called G1L protein) are characterised by a conserved HXXEH motif. MPPs have an active site similar to that of thermolysin and catalyse the removal of N-terminal targeting signals from mitochondrial proteins synthesized in the cytoplasm.

Clan MF, comprising family M17, only contains eukaryotic and bacterial leucyl aminopeptidases both of which require two metal ions for catalytic activity.

A number of very dissimilar exopeptidases belong to clan MG (family M24). Most of these enzymes require two cobalt or manganese ions. The most important members of this family are the bacterial methionyl aminopeptidase and X-Pro aminopeptidase, as well as the type I (mitochondrion) and type II (cytoplasmic) eukaryotic methionyl aminopeptidases that cleave the initial methionine co-translationally from newly synthesized proteins. These enzymes have been shown to comprise a group of proteases occurring ubiquitously in all genomes.

Members of clan MH, which is further divided into families M20, M28 (mainly carboxy- and aminopeptidases), M18 and M42 (mainly aminopeptidases), require co-catalytic zinc ions.

Members of clan MJ (families M19 and M38) are dipeptidases, one of which cleaves rather exotic substrates having isoaspartyl residues.

The only peptidase of clan MK (family M22) is the O-sialoglycoprotein endopeptidase. The other members of this clan are ubiquitously present in all genomes and are characterised by a fold that seems to be similar to that of the non-metalloproteins actin, Hsp70 and DnaK.

The most important feature of the M50 family members of clan MM is the presence of an HEXXH motif like that of clan MA. These enzymes are bound to membranes, contain one zinc ion and have been shown to be involved in the regulation of gene expression by proteolytic processing of transcription regulators.

Members of clan MN (family M55), represented by D-aminopeptidase DppA, contain co-catalytic zinc ions.

Members of clan MO (family M23), such as β-lytic metallopeptidase from *Achromobacter lyticus*, are endopeptidases that lyse bacterial cell wall peptidoglycans.

Clan MP, family M67 comprises isopeptidases (*e.g.* Poh1 peptidase from *Saccharomyces cerevisiae*) that release ubiquitin from ubiquitinated proteins.

Clan MQ, family M29 includes aminopeptidases from thermophilic bacteria such as aminopeptidase T from *Thermus aquaticus* and PepS aminopeptidase from *Streptococcus thermophilus*.

Clan M- contains metallopeptidase families that have not yet been assigned to a well-defined clan. It comprises families M49 (represented by dipeptidyl-peptidase III from *Rattus norvegicus* which releases N-terminal dipeptides sequentially from peptides like angiotensins II and III, Leu-enkephalin, prolactin and alpha-melanocyte-stimulating hormone), M73 (camelysin, a cell-surface endopeptidase from *Bacillus cereus*) and M75 (imelysin or 'insulin-cleaving membrane protease' from *Pseudomonas aeruginosa*).

Thermolysin and thermolysin-like proteases, which are the only metalloproteases that have achieved industrial application, will be discussed in detail in the following sections of this chapter.

2. THERMOLYSIN - THE PROTOTYPE OF METALLOPROTEASES

Thermolysin (EC 3.4.24.27; MEROPS classification: M04.001, family M4 of subclan MA(E)) is an extracellular metalloendoproteinase produced by the gram-positive bacterium *Bacillus thermoproteolyticus* (Endo, 1962). Its three-dimensional structure has been resolved (Matthews *et al*., 1972). Family M4 contains several other extracellular metalloproteases that are produced by various gram-positive bacilli. Phylogenetic analysis (Barrett *et al*., 2004) reveals the occurrence of a group of very closely related enzymes and the existence of others which are less related both in sequence and properties. In the following, only thermolysin and some of the most closely related members of family M4, the thermolysin-like proteases (TLPs), are discussed. The enzymes from *B. brevis*, *B. polymyxa*, *B. megaterium*, *B. amyloliquefaciens*, *B. amylosaccharitic*us and *Lactobacilli* are less well characterised and therefore not considered.

As the alignment in Fig. 1 shows, TLPs from *B. stearothermophilus* (SwissProt P06874; Fujii *et al*., 1983) and *B. caldolyticus* (SwissProt P23384; van den Burg *et al*., 1991) are very similar to thermolysin (SwissProt P00800) whereas the enzymes from *B. subtilis* (SwissProt P39899; Tran *et al*., 1991) and *B. cereus* (UniProt P05806; Wetmore *et al*., 1992) are more distantly related.

3. MOLECULAR STRUCTURE AND FUNCTION

In addition to thermolysin, the crystal structures of other members of clan MA have been resolved (*e.g.* TLP from *B. cereus* - Stark *et al*., 1992, and pseudolysin, the elastase from *Pseudomonas aeruginosa* - Thayer *et al*., 1991). Based on these structures, a homology model of the enzyme from *B. stearothermophilus* has been proposed (Vriend and Eijsink, 1993).

As detected later (Holland *et al*., 1992), the first thermolysin structure published (Matthews *et al*., 1972) contained a dipeptide at the active site. Since then the structure has been refined several times and has now also been resolved both in the presence of inhibitors (Matthews, 1988; Hausrath and Matthews, 2002) and in free form (Hausrath and Matthews, 2002). The mature enzyme consisting of 316 amino acid residues comprises two domains – the N-terminal domain which mainly consists of β-sheets and minor α-helical parts, and the C-terminal domain which is predominantly α-helical in character. The active site cleft with the catalytically essential zinc ion is located between the two domains, and the HEXXH motif (amino acid residues 142 - 146) is integrated into the central α-helix (Fig. 2). The spacing of the zinc ligands follows a short-long pattern as in all members of this clan, *i.e.* the first two ligands are arranged close together (H142 and H146) and the

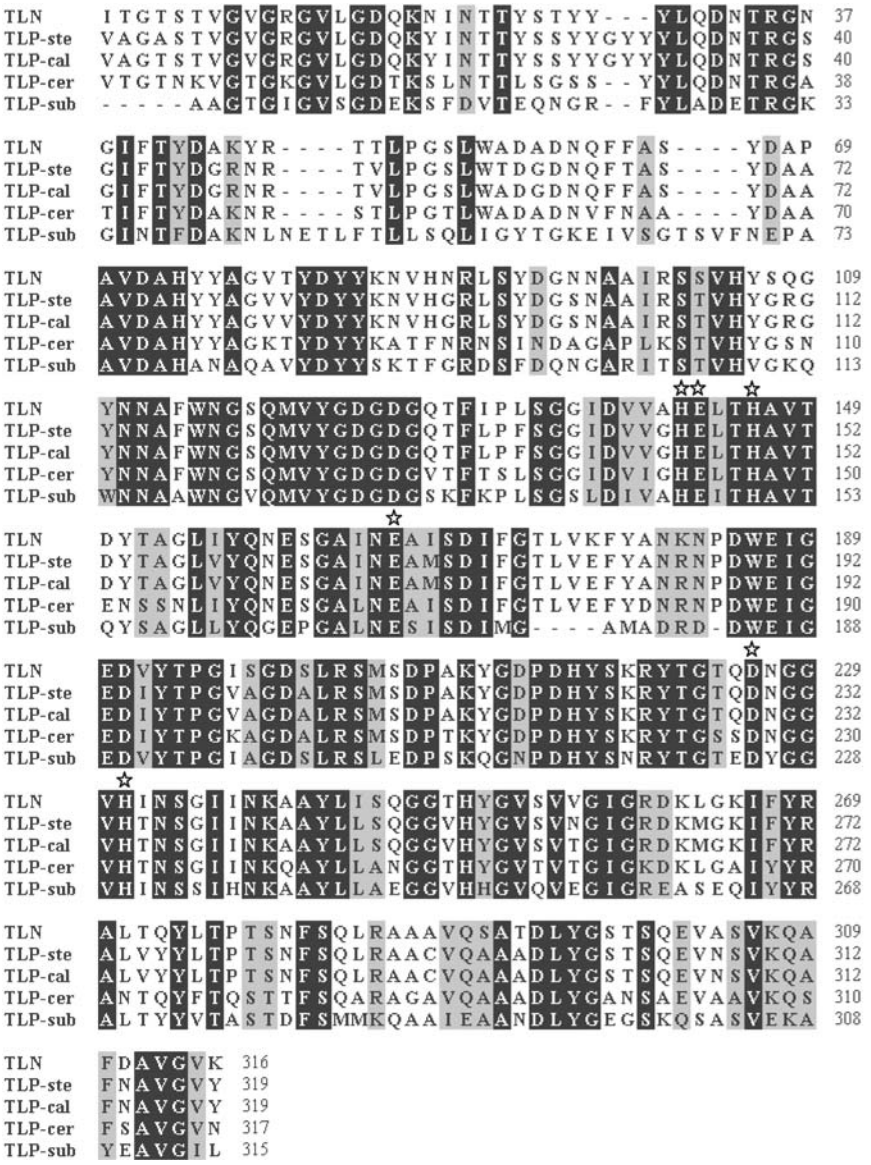


Figure 1. Alignment of the amino acid sequences of selected thermolysin-like proteases. The alignment was created by CLUSTAL W (Thompson *et al*., 1994) in BioEdit 6.0.7. Stars indicate the amino acids of the HEXXH motif and other residues assumed to be involved in catalysis. TLN – thermolysin, TLP-ste – TLP from *B. stearothermophilus*, TLP-cal – TLP from *B. caldolyticus*, TLP-cer – TLP from *B. cereus*, TLP-sub – TLP from *B. subtilis*. Identical amino acids are highlighted by white letters on a black background; similar amino acids are shown in dark grey letters on a light grey background

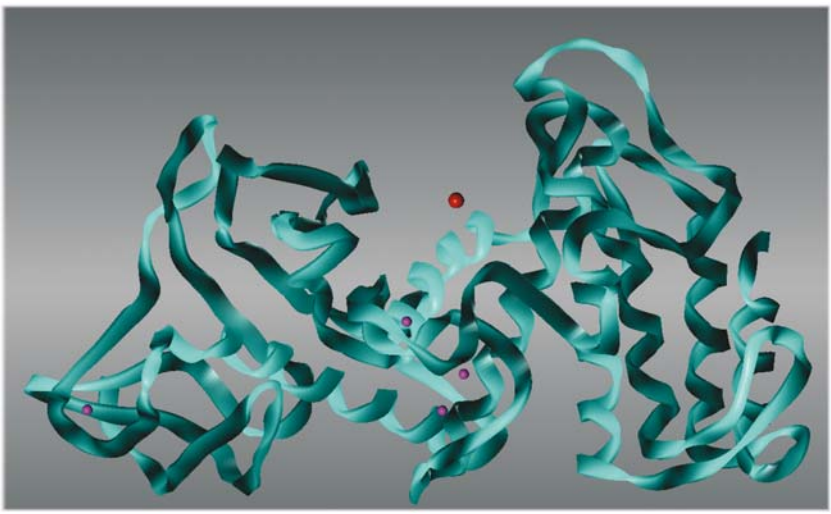


Figure 2. Structure of thermolysin from *B. thermoproteolyticus*. The larger sphere in the centre represents the Zn^{2+} ion in the active site. The four smaller spheres represent the Ca^{2+} ions bound to the molecule

third one (E166) is more distant in the C-terminal direction (Figs. 1 and 3). The fourth ligand represents a water molecule. Removal of the tetrahedrally coordinated zinc ion leads to an inactive enzyme (Holmquist and Valled, 1974) the activity of which towards furylacryloyl-glycyl-L-leucine amide (FAGLA) can be restored by addition of stoichiometric amounts of Zn^{2+}, Co^{2+} and Mn^{2+} ions. Excess of Zn^{2+} ions inhibits the enzyme (Holmquist and Valled, 1974). In the crystal structures of metal-substituted thermolysin derivatives, conformational changes in the active site cleft have been observed (Holland *et al*., 1995).

In addition to the catalytic zinc ion, thermolysin and the other more thermostable TLPs possess four calcium ions which are principle determinants of the stability of these enzymes. The less thermostable TLPs have only two calcium binding sites. Calcium ions 1 and 2 are bound at a double binding site near the active site cleft, whereas calcium ions 3 and 4 are bound to exposed loops in the N-terminal or C-terminal domains, respectively (Matthews *et al*., 1972).

3.1. Catalytic Mechanism

Despite numerous crystallographic, kinetic, inhibition, quantum chemical and site-directed mutagenesis studies, the mechanism of peptide hydrolysis by thermolysin remains controversial. Nevertheless, the zinc ion plays the main role in all the mechanistic proposals. The first mechanism to be proposed (Hangauer *et al*., 1984; Matthews, 1988) was based on the polarisation of the zinc-bound water molecule by the glutamate residue in the HEXXH motif (Glu143 in thermolysin, Fig. 3)

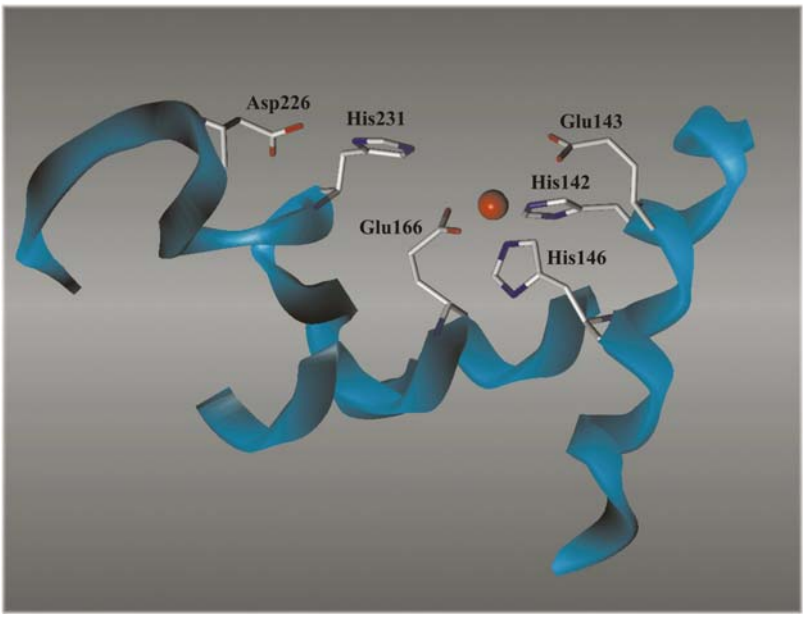


Figure 3. The active site of thermolysin. Amino acids involved in catalysis and mentioned in the text are shown in detail

and its subsequent nucleophilic attack on the carbonyl carbon atom of the scissile bond. A protonated His residue (His231 in thermolysin), which is bound via a hydrogen bond to an Asp residue in the active site cleft (Asp226 in thermolysin), is claimed to stabilise the transition state by forming a hydrogen bond to the carbonyl oxygen atom. The breakdown of the tetrahedral intermediate in the transition state is accomplished by protonation of the amide nitrogen of the scissile bond via Glu143.

An alternative proposal, the 'reverse protonation' mechanism, derives from the observed dependence of the catalytic constants on pH in the hydrolysis of arazoformyl peptides which are poor substrates for thermolysin. This alternative questions the role of Glu143 as a general base and instead proposes the assignment of this function to a His residue (His231 in thermolysin) in the active site (Mock and Aksamawati, 1994; Mock and Stanford, 1996). In this case, the scissile bond is activated for hydrolysis by direct coordination of the carbonyl oxygen atom to the zinc ion as a potent Lewis acid with simultaneous displacement of the water molecule otherwise bound to the zinc ion. By analogy to serine proteases, His231 along with Asp226 is supposed to enable proton transfer by deprotonation of an incoming water molecule which is not bound to the zinc ion.

A recent quantum chemical study (Pelmenschikov *et al*, 2002) has confirmed the key role of Glu143 in the thermolysin-catalysed hydrolysis of peptides, thus supporting the initially proposed mechanism of Matthews and co-workers. In the context of these data, the reverse protonation mechanism seems to be less favoured.

Glu143 is absolutely conserved in all metalloproteases having the HEXXH motif, whereas His231 is only partly conserved in metalloproteases employing the same catalytic mechanism. Additional arguments strengthening the essential role of Glu143 are provided by site-directed mutagenesis studies. The charge-conserving replacement of Glu143 by Asp in NprM from *B. stearothermophilus* MK232, an enzyme identical to thermolysin at the primary structure level, led to complete loss of activity (Kubo *et al*., 1992). This makes it rather unlikely that Glu143 acts solely as a negatively charged counter-ion providing electrostatic stabilisation of the transition state. Replacement of Glu143 in TLP from *B. subtilis* caused nearly complete loss of secreted enzyme, whereas His231 (thermolysin numbering) mutants were secreted and retained a certain degree of activity (Toma *et al*., 1989). Replacement of His231 by Ala or Phe in the TLP from *B. stearothermophilus* reduced the k_{cat}/K_m value 430 and 500-fold, respectively (Beaumont *et al*., 1995). These mutants showed reduced pH dependence in the alkaline range. Bearing all this in mind, the essential role of His231 is not supported. Attempts performed by our group to produce completely inactive TLPs from *B. stearothermophilus* showed that the least active enzyme was the Glu143Gln mutant having less than 0.1% residual activity, whereas the His231Ala mutant enzyme restored 1.6% of wild-type activity. A combination of both mutations was not able to decrease the activity further (Mansfeld, unpublished results). The enzymes were expressed in *E. coli* and renatured as described in Mansfeld *et al*. (2005).

Based on structural comparisons, a hinge-bending motion leading to the closure of the active site cleft upon substrate or inhibitor binding has been proposed (Holland *et al*., 1992, 1995; Hausrath and Matthews, 2002). The involvement of conserved glycine residues at positions 78, 135 and 136 has been assumed and experimentally proven for thermolysin (Holland *et al*., 1992), a TLP from *B. cereus* (Stark *et al*., 1992) and a TLP from *B. stearothermophilus* (Veltman *et al.*, 1998). Recent comparisons of the 'open' and the 'closed' structures (Hausrath and Matthews, 2002) have however shown that these two regions cannot account completely for the observed movement of the domains at the active site cleft. The concerted movement of a group of side chains is proposed instead.

Thermolysin and related proteases preferably cleave substrates having bulky hydrophobic residues (Leu, Phe) in the P_1' position and smaller amino acids in position P_1 (nomenclature according to Schechter and Berger, 1967). Four major substrate binding pockets (S_2, S_1, S_1', S_2') on the enzyme have been identified (Hangauer *et al*., 1984). The hydrophobic substrate binding pocket S_1' of thermolysin, mainly formed by Phe130, Leu133, Val139 and Leu202, is considered to be the main determinant of substrate specificity and preferably binds hydrophobic residues such as Leu (*e.g.* Hangauer *et al*., 1984; Matthews, 1988). Mutagenesis studies on the TLP of *B. stearothermophilus* have shown that the preference of this protease for Phe at P_1' can be altered toward that of thermolysin by changing Phe133 to Leu, the latter residue being present in thermolysin at that position (de Kreij *et al.*, 2000, 2001). Enlargement of the binding pocket by replacement of Leu202 by smaller amino acids (Val, Ala, Gly) resulted in higher efficiency toward

substrates with Phe at P_1'. Unexpectedly, reduction of its size by substitution of Leu202 by Phe or Tyr also caused a large increase in activity toward substrates with Phe at P_1'.

The role of other substrate binding pockets has recently been highlighted by the adaptation of vimelysin (*Vibrio* sp.) substrate specificity in the P_3' position to that of thermolysin upon exchange of Arg215 in the S_3' binding pocket for Asp215 present in thermolysin (Oda *et al.*, 2005). Vimelysin has 35% sequence identity with thermolysin and is characterised by high stability in organic solvents. It shows a strong preference for Phe over Leu at position P_1' and also a preference for neutral or acidic amino acids in the P_3' position in contrast to thermolysin in which basic amino acids are preferred at position C.

Enhancements of the activity of thermolysin have been achieved by mutation of Tyr110 and Phe114 in the S_2 subsite, as well as Tyr211 (Kubo *et al.*, 1992), Gln119 (Kidokoro *et al.*, 1995) and Leu155 (Matsumiya *et al.*, 2004). Some of the effects were additive (Kidokoro, 1998; de Kreij *et al.*, 2002). However, the effects of mutations on activity were in most cases smaller for large proteinaceous substrates compared to those observed for short peptides probably as a consequence of the different possible productive binding modes for the former. The contribution of binding to catalysis is expected to be much less in the case of short peptides.

Thermolysin activity is considerably enhanced by the addition of neutral salts (Inouye, 1992; Inouye *et al.*, 1998a; Bedell *et al.*, 1998). A further increase in activity is observed when the catalytic zinc ion is substituted by cobalt ions (Kuzuya and Inouye, 2001). Both effects are independent of each other. Activation by NaCl has been shown to be caused by an increase in k_{cat} values (Inouye *et al.*, 1996). The addition of neutral salts also has beneficial effects on thermolysin solubility and thermal stability (Inouye *et al.*, 1998a and b). Preliminary studies of crystals soaked in 4 M NaCl did not show significant changes in the space group (Kamo *et al.*, 2005). Positive effects on thermolysin activity have also been described for sugars (*e.g.* sucrose, trehalose) and other polyols (Mejri *et al.*, 1998), and pressure up to 2.5 kbar (Kudryashova *et al.*, 1998).

Thermolysin is inhibited by zinc-chelating agents (*e.g.* 1,10-phenanthroline, EDTA) (Holmquist and Vallee, 1974). Other more specific inhibitors for thermolysin have been developed over the years, and their number is still increasing (reviewed in van den Burg and Eijsink, 2004).

3.2. Stability

Apart from high activity, the stability of an enzyme at higher temperatures and toward other denaturing influences is one of the most important criteria for the application of enzymes in biocatalysis. To meet the demands of industry, high operational stability of enzymes is required as this significantly contributes to cost reduction.

In order to obtain highly stable enzymes several strategies are used. One is focused on the isolation of new enzymes from extremophilic organisms (reviewed

by Vieille and Zeikus, 2001). Other strategies are based on: the stabilisation of available enzymes by rational design of enzyme variants (Eijsink *et al.*, 2004), high-throughput screening of randomly generated mutant libraries (Eijsink *et al.*, 2005) in combination with recently described 'semi-rational' approaches for a guided design of these libraries (Patrick and Firth, 2005), chemical modification including immobilisation (Ulbrich-Hofmann *et al.*, 1999) and *de novo* design of catalysts (Kaplan and DeGradd, 2004).

Since thermolysin is produced by a highly thermophilic *Bacillus* strain it is relatively thermostable compared to other metalloproteases produced by less thermophilic strains. As a result of a series of mutational studies a surface-exposed region between amino acid residues 56 – 69 in the N-terminal part of the thermolysin-like protease from *B. stearothermophilus* (TLP-ste) (Takagi *et al.*, 1985; 85% sequence identity to thermolysin) has been identified that is extremely sensitive to mutation, whereas the C-terminal part of the protease is only slightly affected by even dramatic amino acid changes (Vriend and Eijsink, 1993; Eijsink *et al.*, 1995). Later this region was recognised as being the most labile region of the protein where local unfolding processes start, resulting in rapid autoproteolytic degradation of the protein (Eijsink *et al.*, 1995; Vriend *et al.*, 1998; Fig. 4). In light of the concept of protein stabilisation developed by Schellenberger and Ulbrich (1989), this region was consequently called the unfolding region (Mansfeld *et al.*, 1999). Coherently, stabilisation of this region enabled the construction of variants displaying considerably enhanced thermostability. The amino acid residues

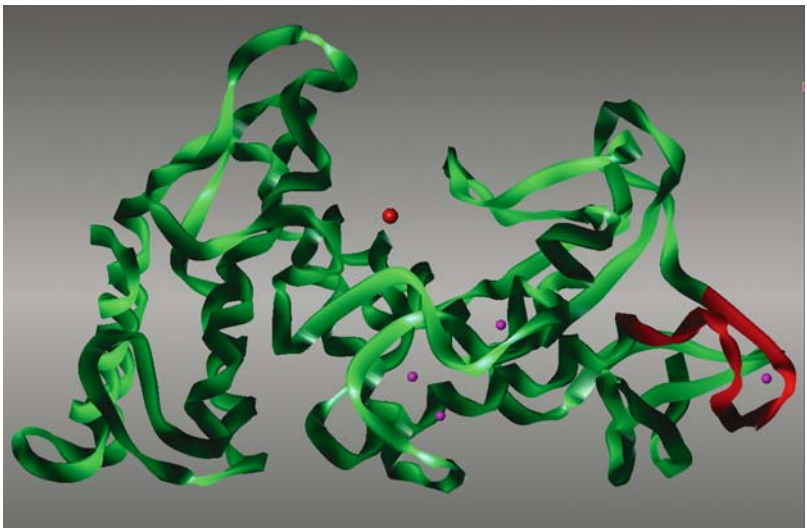


Figure 4. Homology model of the 3D structure of TLP from *B. stearothermophilus* (Vriend and Eijsink, 1993). The large sphere represents the Zn^{2+} ion in the active site. The four smaller spheres represent the Ca^{2+} ions bound to the molecule

selected for replacement were chosen on the one hand on the basis of rational design strategies, and on the other because they corresponded to residues naturally occurring in the more thermostable enzyme thermolysin (reviewed in Eijsink *et al.*, 1995). The introduction of a disulfide bridge between residues 8 and 60 (Mansfeld *et al*, 1997) was found to strongly stabilise this region against local unfolding. Detailed studies of this mutant have shown that this effect is mainly attributable to stabilisation against autoproteolysis rather than global unfolding (Dürrschmidt *et al*, 2001). The disulfide bridge was shown to be able to mimic the stabilising effect of calcium ions in local unfolding processes (Dürrschmidt *et al.*, 2005). Calcium ions are major determinants of protease stability (Dahlquist *et al.*, 1976).

These studies culminated in the successful conversion of the moderately stable TLP from *B. stearothermophilus* into an extremely stable enzyme (named boilysin) via a limited number of mutations (van den Burg *et al*, 1998). The half-life of the mutant enzyme was 170 min at 100°C in contrast to 1 min for thermolysin. Compared to naturally occurring enzymes from thermophilic organisms, boilysin is characterised by wild-type like activity under the usually employed operating temperatures making this enzyme interesting for industrial application. As temperature is increased, activity is further enhanced. This enzyme has been tested under extreme conditions for the hydrolysis of substrates that are difficult to digest (van den Burg *et al.*, 1999; de Kreij *et al*, 2000), the hydrolysis of prion proteins and protein removal in nucleic acid purification (van den Burg, personal communication), and for the synthesis of an aspartame precursor (Kühn *et al*, 2002).

Taking advantage of knowledge on the enzyme's unfolding process and the concept of stabilisation by strengthening the most labile region in a protein (Schellenberger and Ulbrich, 1989; Ulbrich-Hofmann *et al*, 1999), strong stabilisation of TLP-ste was achieved by immobilisation in a site-specific manner (Mansfeld *et al.*, 1999). This was most effective when the protein was fixed to the carrier via cysteine residues in the unfolding region. Very strong stabilisation has also been obtained by immobilisation via multiple bonds to a carrier having a high density of functional groups (Mansfeld and Ulbrich-Hofmann, 2000), suggesting that more than one labile region might be present in the molecule as has been argued by Vriend *et al.* (1998). Rigidification of the enzyme due to the formation of multiple bonds with the carrier material was found to occur at the expense of activity (Mansfeld and Ulbrich-Hofmann, 2000).

Another strategy to protect TLPs against autoproteolytic degradation and to stabilise against thermal inactivation is the identification of primary cleavage sites. The removal of these autodegradation sites has been described for TLP from *B. subtilis* (van den Burg *et al*, 1990) and thermolysin (Matsumiya *et al*, 2004).

4. APPLICATIONS

Thermolysin is commercially available at industrial scale as Thermoase PC10F from Amano Enzyme Inc., formerly Daiwa Kasei Co. Ltd., Japan. Other bacterial metalloproteases produced commercially are the TLPs from *B. subtilis* (Neutrase® from

Novo Nordisk, Denmark and Protin PC10F from Amano Enzyme Inc., Japan). The highly stable TLP-ste variant Boilysin can be requested from IMEnz (Groningen, The Netherlands).

At laboratory scale, thermolysin can be purchased from different suppliers (*e.g.* Sigma, MERCK Biosciences). Thermolysin and TLPs from *Bacillus* species are traditionally produced in protease-deficient *B. subtilis* strains such as DB104 (Kawamura and Doi, 1984) and DB117 (Eijsink *et al.*, 1990). These enzymes are synthesized as inactive preproenzymes (Takagi *et al.*, 1985; Kubo and Imanaka, 1988; van den Burg *et al.*, 1991) and processed to the active mature enzymes via autocatalytic removal of the large propeptides (about 200 amino acids) (Wetmore *et al.*, 1992; Marie-Claire *et al.*, 1998). The thermolysin propeptide has been shown to act as a mixed, non-competitive inhibitor of the protease and to facilitate the recovery of active enzyme from denatured thermolysin in a stoichiometric manner (O'Donohue and Beaumont, 1996; Marie-Claire *et al.*, 1999). Recently, several strategies have been described for the expression of these enzymes in *E. coli* with the prosequence in *cis* or *trans* (Marie-Claire *et al.*, 1999; Inouye *et al.*, 2005) and even in the absence of the prosequence (Mansfeld *et al.*, 2005).

The enzymes secreted into the culture broths of *B. subtilis* and *E. coli* or renatured from inclusion bodies formed in *E. coli* (Mansfeld *et al.*, 2005) can be purified by affinity chromatography on Bacitracin-silica (van den Burg *et al.*, 1989) or Gly-D-Phe columns (Walsh *et al.*, 1974).

4.1. Synthesis of Peptides

Large scale synthesis of peptides has become increasingly important for the food and pharmaceutical industries over the last few decades. The main application of peptides is their use as low-calorie sweeteners. In addition, several biologically active peptides have found interest as drugs in the treatment of diseases. Apart from conventional peptide synthesis, new strategies for production have been tested, one of which is enzymatic synthesis. The advantages of using enzymes are the stereospecificity they confer on the reaction, the necessity for only minimal side chain protection, the mild reaction conditions, and the avoidance of racemisation.

Thermolysin is one of the enzymes that has been studied for its potential in enzymatic peptide synthesis. Since its first use for this purpose (Isowa *et al.*, 1979) it has been extensively studied in the laboratory and is used for large-scale production of N-carbobenzoxy-L-aspartyl-L-phenylalanine methyl ester (Z-Asp-Phe-OMe), the precursor of the widely used artificial sweetener aspartame (Ooshima *et al.*, 1985; Murakami *et al.*, 1996). Thermolysin proved to be advantageous in the synthesis of aspartame due to its low esterolytic activity that results in the preservation of the methyl ester group which is essential for the sweet taste of the peptide.

As no acyl enzyme is formed in thermolysin catalysis, the reactions cannot be run in kinetically controlled mode. Therefore, several strategies have been tested to shift the equilibrium of the thermodynamically controlled enzymatic synthesis of Z-Asp-Phe-OMe to the desired product. These are based on aqueous systems

(Inouye, 1992; Murakami *et al*., 1996), systems with water-miscible organic solvents (Lee *et al*., 1992; Kühn *et al*., 2002), biphasic systems (Hirata *et al*., 1997; Murakami and Hirata, 1997; Murakami *et al*., 1998; Miyanaga *et al*., 2000b), solid-to-solid synthesis (Erbeldinger *et al*., 1998a, b; Erbeldinger *et al*., 2001) and low-water solvent systems (Nakanishi *et al*., 1985). For syntheses in aqueous/organic biphasic systems (Murakami and Hirata, 1997; Hirata *et al*., 1997), in low-water solvent systems with immobilized thermolysin (Nakanishi *et al*., 1985, 1990) or in membrane systems (Iacobucci *et al*., 1994), continuous operation has been used successfully.

Yields in pure aqueous systems are usually very low. The activity of thermolysin and, accordingly, the reaction rates in aqueous systems have been found to be enhanced by the addition of sodium and potassium salts (Inouye, 1992). However, the pH increase due to salt addition may result in non-enzymatic hydrolysis of the reactants, a problem which is avoided in reactions at low pH. High yields (95%) have been achieved by insoluble salt formation between the product and excess Phe-OMe or unreacted enantiomer D-Phe-OMe followed by subsequent removal of the precipitate from the aqueous solution (Ager *et al*., 1998). This method is used in the commercial aspartame precursor production process of TOSOH (Japan) performed at Holland Sweetener (The Netherlands) (Fig. 5). Addition of a water-immiscible solvent like toluene or 4-methylpentan-2-one after the start of formation of the precipitate was found to permit the process to be run continuously. In the presence of water-miscible organic solvents (*e.g.* dimethylsulfoxide) reaction rates were also enhanced by the addition of salts (Kühn *et al*., 2002), though yields decreased with increasing salt concentrations. Yields could be markedly improved by the addition of alcohols (methanol, 2-propanol) to aqueous systems even though reaction rates were reduced due to inhibitory effects on thermolysin (Kühn *et al*., 2002). In biphasic organic solvent systems, ethyl acetate, tert.-amyl alcohol (Miyanaga *et al*., 2000b), n-butyl acetate (Murakami and Hirata, 1997), tributylphosphate and 1-butanol (in the synthesis of N-formyl-Asp-Phe-OMe – Murakami *et al*., 2000a) and ionic liquids (*e.g.* 1-butyl-3-methylimidazolium hexafluorophosphate – Erbeldinger *et al*., 2000) have all been used as solvents. In the solid-to-solid system the pH adjusted by basic inorganic salt addition played an important role (Erbeldinger *et al*., 2001). In low-water solvent systems the water is usually provided by the carrier materials that are used for adsorptive binding of the enzyme. Polyacrylic ester resins such as XAD-7 (ICN Biomedicals Inc., USA) in ethyl acetate and tert.-amyl alcohol (Miyanaga *et al.*, 2000a, b), Celite R-640 (FLUKA) in combination with toluene as solvent (de Martin *et al*., 2001) or molecularly imprinted polymers (methacrylate-ethylene glycol dimethacrylate-copolymers) in ethyl acetate (Ye *et al*., 1999) have all been used as carrier materials. A considerable increase in thermolysin activity in non-aqueous media has been achieved by lyophilisation in the presence of KCl or other inorganic salts (Bedell *et al*., 1998). Activity could be further improved by the use of molecular imprinting in combination with activation by salts (Rich *et al*., 2002). Cross-linked enzyme crystals (CLECs) of thermolysin which have been used

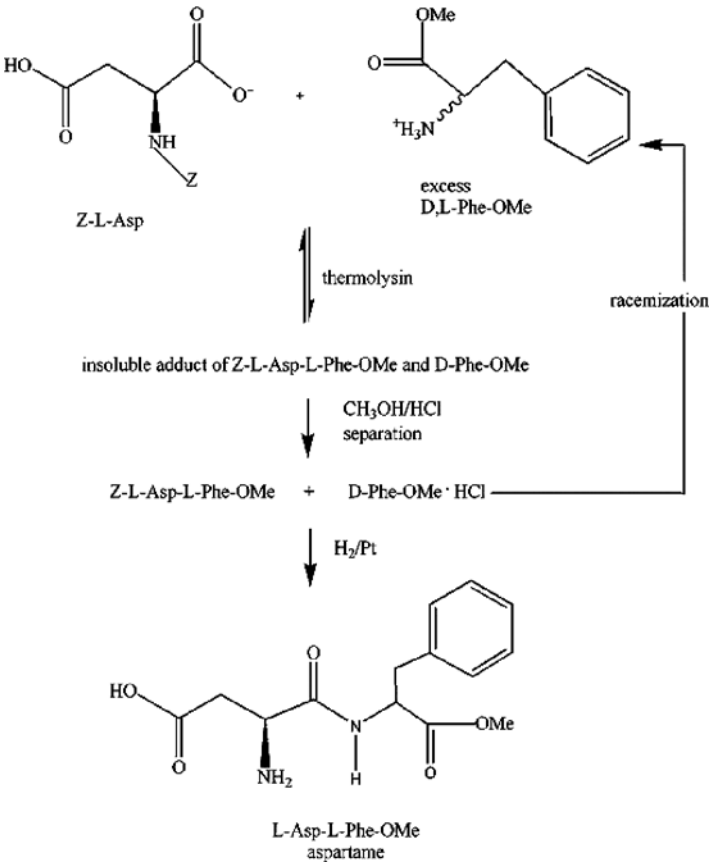


Figure 5. Principle of commercial aspartame synthesis of TOSOH at Holland Sweetener (The Netherlands)

successfully for the synthesis of the aspartame precursor (Persichetti *et al.*, 1995) represent an interesting tool for organic chemists due to their high specific activity and increased resistance to inactivation by organic solvents, elevated temperatures and proteolysis.

In all these systems the reaction velocities and yields obtained result from the interplay between the reaction medium (*i.e.* buffer, organic solvent, salt concentration), the types and ratios of reactants, and the type of product, as well as the activity and stability of thermolysin in the corresponding reaction systems. A compromise always has to be found between initial reaction rates and final yields.

To broaden the scope for enzymatic peptide synthesis new enzymes have been searched for. A metalloprotease called vimelysin from *Vibrio* sp. proved to be superior to thermolysin for the synthesis of aspartame at lower temperatures and higher solvent concentrations. Like vibriolysin from the Antarctic bacterium

strain 643 (Adekoya *et al*., 2006), it might be an interesting alternative for thermolabile substrates (Kunugi *et al*., 1997). Due to its broader substrate specificity, pseudolysin from *Pseudomonas aeruginosa* might also be interesting for synthetic purposes (Rival *et al*., 2000). At laboratory scale, free or immobilized thermolysin has also been used for the synthesis of other peptides: Z-Gln-Leu-NH_2, Z-Phe-Leu-NH_2, various dipeptide fragments of cholecystokinin, and peptides containing non-proteogenic amino acids (Wayne and Fruton, 1983; Erbeldinger *et al.*, 1998a, b, 1999; Calvet *et al*., 1996; Krix *et al*., 1997). Neutrase®, a TLP from *B. subtilis*, with a slightly different substrate specificity to that of thermolysin, was also applied in peptide synthesis (either as free or immobilized enzyme (on Celite-545 (Fluka, Germany) or Polyamide-PA6 (Akzo)) (Clapes *et al*., 1995, 1997). A scale-up of the suspension-to-suspension approach using mainly undissolved substrates was performed by Eichhorn *et al*. (1997). The possibility of using the extractive reaction in aqueous/organic biphasic systems for the continuous synthesis of Z-Gly-Phe-OMe at larger scale and in high yield was reported by Murakami *et al.* (2000b). Unexpectedly, thermolysin was also shown to catalyse the acylation of paclitaxel with divinyl adipate (Khmelnitsky *et al*., 1997) and an acylation of sucrose with vinyl laurate (Pedersen *et al*., 2002).

4.2. Production of Protein Hydrolysates

Another very important industrial application of metalloproteases (mostly in combination with other proteases) is the production of hydrolysed food proteins and flavour-enhancing peptides to replace the chemical methods of synthesis. Soy and wheat hydrolysates are used in flavour-enhancement of soups and sauces, and milk protein hydrolysates are preferred for the refinement of cheese products. Meat hydrolysates find application in the enhancement of the flavour of meat products, soups, sauces and other instant products. An advantage of enzymatic processes is the minor formation of unwanted by-products of negative impact on health; the disadvantageous formation of bitter tasting peptides in enzymatically produced protein hydrolysates can be overcome by simultaneous treatment with exopeptidases (reviews in Saha and Hayashi, 2001; Raksakulthai and Haard, 2003). Bitter peptides are characterised by a high content of hydrophobic amino acids. Hydrolysates of meat and fish proteins or gelatin develop less bitter taste than hydrolysates of maize protein, casein or haemoglobin. Flavour development by a cocktail of proteases, including Neutrase® has been used to accelerate the ripening of dry fermented sausages (*e.g.* Fernandez *et al*., 2000). The development of high-value functional foods and nutraceuticals has made a major impact on dairy protein hydrolysate production because of the latter's probiotic, antimicrobial and digestive effects. Another beneficial effect of protein hydrolysates on health might be their antioxidant effects (Hernandez-Ledesma *et al*., 2005) and their inhibition of angiotensin-converting enzyme (Vercruysse *et al*., 2005). Low-molecular weight hydrolysis products of protamine have been tested successfully as a delivery system for DNA in gene therapy (Park *et al*., 2003).

4.3. Other Applications

Further commercial applications of metalloproteases can be found in the brewing industry (improved filtration of beer, reduced calorie content), in the leather industry (bating and dehairing), the processing of slaughter waste, improvement of the baking characteristics of flour, and in the film industry (recovery of waste silver). An interesting application of immobilized thermolysin is the removal of protein coatings from the surface of old documents and art work (Moeschel *et al.*, 2003).

In protein science, thermolysin is an important tool for limited proteolysis to determine primary structures and gain first insights into the conformation of proteins whose crystal structures are not yet known (reviewed in Fontana *et al.*, 2004). It is also used to analyse confined local fluctuations and global unfolding events in proteins and to determine their stabilities (Arnold *et al.*, 2005; Park and Marqusee, 2005), or to isolate protein fragments that can fold autonomously and therefore be considered as domains which might be useful in crystallisation and high-throughput applications (Gao *et al.*, 2005).

Inhibitors of metalloproteases involved in diseases are of potential therapeutic use. In this respect, thermolysin has been used as a template for the creation of homology models of the active sites of medically relevant mammalian metalloproteases. Details of these important classes of metalloproteases can be found in Barrett *et al.* (2004). Interesting targets are: neprilysin (Roques *et al.*, 1993); angiotensin-converting enzyme, being responsible for degradation of biologically active peptides such as enkephalins; endothelin-converting enzymes, which liberate endothelin (a potent vasoconstrictor) from its precursor; highly potent neurotoxins like bontoxilysin and tentoxilysin from *Clostridium* species which block the release of acetylcholine at neuromuscular junctions and cause motor paralysis in tetanus and botulism; anthrax lethal factor, which acts by disrupting intracellular signalling by cleaving mitogen-activated protein kinase kinases and causes multiple haemorrhagic lesions; matrix metalloproteases or matrixins (like collagenase, elastase, stromelysin, matrilysin, gelatinase) which are involved in the degradation of extracellular matrix proteins including collagen and are required for tissue repair and remodelling but are also involved in pathological processes (arthritis, atherosclerosis, tumour growth and metastasis); ADAM17 (tumour necrosis factor α-converting enzyme) and ADAM10 (myelin-associated metalloendopeptidase), pappalysins 1 and 2 which cleave insulin-like growth factor 1 binding protein-4 and liberate the growth factor.

5. ACKNOWLEDGEMENT

I would like to thank Prof. Dr. R. Ulbrich-Hofmann for critical reading of the manuscript, Dr. S. Gebauer's group of Molecular Modelling at our department for help in preparing the Figures, Prof. G. Vriend, Dr. B. van den Burg, Prof. V. Eijsink, and Prof. G. Venema for getting me started with the metalloprotease work and fruitful cooperation.

REFERENCES

Adekoya, O.A., Helland, R., Willassen, N.-P. and Sylte, I. (2006) Comparative sequence and structure analysis reveal features of cold adaptation of an enzyme in the thermolysin family. Proteins *62*, 435–449.

Ager, D.J., Pantaleone, D.P., Henderson, S.A., Katritzky, A.R., Prakash, I. and Walters, D.E. (1998) Commercial, synthetic nonnutritive sweeteners. Angew. Chem. Int. Ed. *37*, 1802–1817.

Arnold, U., Köditz, J., Markert, Y. and Ulbrich-Hofmann, R. (2005) Local fluctuations vs. global unfolding of proteins investigated by limited proteolysis. Biocatal. Biotrans. *23*, 159–167.

Barrett, A.J., Rawlings, N.D. and Woessner, J.F. (Editors) (2004) Handbook of proteolytic enzymes. 2nd edition. Volume 1, pp 231. Elsevier Academic Press, Amsterdam.

Beaumont, A., O'Donohue, M.J., Paredes, N., Rousselet, N., Assicot, M., Bohuon, C., Fournie-Zaluski, M.-C. and Roques, B.P. (1995) The role of histidine 231 in thermolysin-like enzymes. J. Biol. Chem. *270*, 16803–16808.

Bedell, B.A., Mozhaev, V.V., Clark, D.S. and Dordick, J.S. (1998) Testing for diffusion limitations in salt-activated enzyme catalysts operating in organic solvents. Biotechnol. Bioeng. *58*, 654–657.

Bode, W., Gomis-Rüth, F.-X., Huber, R., Zwilling, R. and Stöcker, W. (1992) Structure of astacin and implications for activation of astacins and zinc-ligation of collagenases. Nature *358*, 164–167.

Calvet, S., Torres, J.L. and Clapes, P. (1996) Enzymatic peptide synthesis in organic media. Synthesis of CCK-8 dipeptide fragments. Biocatal. Biotrans. *13*, 201–216.

Clapes, P., Pera, E. and Torres, J.L. (1997) Peptide bond formation by the industrial protease, neutrase, in organic media. Biotechnol. Lett. *19*, 1023–1026.

Clapes, P., Torres, J.L. and Adlercreutz, P. (1995) Enzymatic peptide synthesis in low water content systems: Preparative enzymatic synthesis of [Leu]- and [Met]-enkephalin derivatives. Bioorg. Med. Chem. *3*, 245–255.

Dahlquist, F.W., Long, J.W. and Bigbee, W.L. (1976) Role of calcium in thermal stability of thermolysin. Biochemistry *15*, 1103–1111.

de Kreij, A., van den Burg, B., Veltman, O.R., Vriend, G., Venema, G. and Eijsink, V.G.H. (2001) The effect of changing the hydrophobic S_1' subsite of thermolysin-like proteases on substrate specificity. Eur. J. Biochem. *268*, 4985–4991.

de Kreij, A., van den Burg, B., Venema, G., Vriend, G., Eijsink, V.G.H. and Nielsen, J.E. (2002) The effects of modifying the surface charge on the catalytic activity of a thermolysin-like protease. J. Biol. Chem. *277*, 15432–15438.

de Kreij, A., Venema, G. and van den Burg, B. (2000) Substrate specificity in the highly heterogeneous M4 peptidase family is determined by a small subset of amino acids. J. Biol. Chem. *275*, 31115–31120.

de Martin, L., Ebert, C., Gardossi, L. and Linda, P. (2001) High isolated yields in thermolysin-catalysed synthesis of Z-L-aspartyl-L-phenylalanine methyl ester in toluene at controlled water activity. Tetrahedron Lett. *42*, 3395–3397.

Dürrschmidt, P., Mansfeld, J. and Ulbrich-Hofmann, R. (2001) Differentiation between conformational and autoproteolytic stability of the neutral protease from *Bacillus stearothermophilus* containing an engineered disulfide bond. Eur. J. Biochem. *268*, 3612–3618.

Dürrschmidt, P., Mansfeld, J. and Ulbrich-Hofmann, R. (2005) An engineered disulfide bridge mimics the effect of calcium to protect neutral protease against local unfolding. FEBS J. *272*, 1523–1534.

Eichhorn, U., Bommarius, A.S., Drauz, K. and Jakubke, H.-D. (1997) Synthesis of dipeptides by suspension-to-suspension conversion via thermolysin catalysis: from analytical to preparative scale. J. Peptide Sci. *3*, 245–251.

Eijsink, V.G.H., Bjork, A., Gaseidnes, S., Sirevag, R., Synstad, B., van den Burg, B. and Vriend, G. (2004) Rational engineering of enzyme stability. J. Biotechnol. *113*, 105–120.

Eijsink, V.G.H., Gaseidnes, S., Borchert, T.V. and van den Burg, B. (2005) Directed evolution of enzyme stability. Biomol. Eng. *22*, 21–30.

Eijsink, V.G.H., Veltman, O.R., Aukema, W., Vriend, G. and Venema, G. (1995) Structural determinants of the stability of thermolysin-like proteinases. Nature Struct. Biol. *2*, 374–379.

Eijsink, V.G.H., Vriend, G., van den Burg, B., Venema, G. and Stulp, B.K. (1990) Contribution of the C-terminal amino acid to the stability of *Bacillus subtilis* neutral protease. Protein Eng. *4*, 99–104.

Endo, S. (1962) The protease produced by thermophilic bacteria. Hakko Kogaku Zashi *40*, 346–347.

Erbeldinger, M., Mesiano, A.J. and Russell, A.J. (2000) Enzymatic catalysis of formation of Z-aspartame in ionic liquid - An alternative to enzymatic catalysis in organic solvents. Biotechnol. Progr. *16*, 1129–1131.

Erbeldinger, M., Ni, X. and Halling, P.J. (1998a) Effect of water and enzyme concentration on thermolysin-catalyzed solid-to-solid peptide synthesis. Biotechnol. Bioeng. *59*, 68–72.

Erbeldinger, M., Ni, X. and Halling, P.J. (1998b) Enzymatic synthesis with mainly undissolved substrates at very high concentrations. Enzyme Microb. Technol. *23*, 141–148.

Erbeldinger, M., Ni, X. and Halling, P.J. (1999) Kinetics of enzymatic solid-to-solid peptide synthesis: intersubstrate compound, substrate ratio, and mixing effects. Biotechnol. Bioeng. *63*, 316–321.

Erbeldinger, M., Ni, X. and Halling, P.J. (2001) Kinetics of enzymatic solid-to-solid peptide synthesis: synthesis of Z-aspartame and control of acid-base conditions by using inorganic salts. Biotechnol. Bioeng. *72*, 69–76.

Fernandez, M., Ordonez, J.A., Bruna, J.M., Herranz, B. and de la Hoz, L. (2000) Accelerated ripening of dry fermented sausages. Trends Food Sci. Technol. *11*, 201–209.

Fontana, A., de Laureto, P.P., Spolaore, B., Frare, E., Picotti, P. and Zambonin, M. (2004) Probing protein structure by limited proteolysis. Acta Biochim. Pol. *51*, 299–321.

Fujii, M., Takagi, M., Imanaka, T. and Aiba, S. (1983) Molecular cloning of a thermostable neutral protease gene from *Bacillus stearothermophilus* in a vector plasmid and its expression in *Bacillus stearothermophilus* and *Bacillus subtilis*. J. Bacteriol. *154*, 831–837.

Gao, X., Bain, K., Bonnano, J.B., Buchanan, M., Henderson, D., Lorimer, D., Marsh, C., Reynes, J.A., Sauder, J.M., Schwinn, K., Thai, C. and Burley, S.K. (2005) High-throughput limited proteolysis/mass spectrometry for protein domain elucidation. J. Struct. Funct. Genomics *6*, 129–134.

Hangauer, D.G., Monzingo, A.F. and Matthews, B.W. (1984) An interactive computer graphics study of thermolysin-catalyzed peptide cleavage and inhibition by N-carboxymethyl dipeptides. Biochemistry *23*, 5730–5741.

Hausrath, A.C. and Matthews, B.W. (2002) Thermolysin in the absence of substrate has an open conformation. Acta Crystallogr. D, Biol. Crystallogr. *58*, 1002–1007.

Hernandez-Ledesma, B., Davalos, A., Bartolome, B. and Amigo, L. (2005) Preparation of antioxidant enzymatic hydrolysates from α-lactalbumin and β-lactoglobulin. Identification of active peptides by HPLC-MS/MS. J. Agric. Food Chem. *53*, 588–593.

Hirata, M., Ishimine, T. and Hirata, A. (1997) Development of novel method for enzymatic peptide synthesis utilizing extractive reaction. J. Chem. Eng. Jpn. *30*, 467–477.

Holland, D.R., Hausrath, A.C., Juers, D. and Matthews, B.W. (1995) Structural analysis of zinc substitutions in the active site of thermolysin. Protein Sci. *4*, 1955–1965.

Holland, D.R., Tronrud, D.E., Pley, H.W., Flaherty, K.M., Stark, W., Jansonius, J.N., McKay, D.B. and Mattews, B.W. (1992) Structural comparison suggests that thermolysin and related enzymes undergo hinge-bending motion during catalysis. Biochemistry *31*, 11310–11316.

Holmquist, B. and Vallee, B.L. (1974) Metal substitutions and inhibition of thermolysin: spectra of the cobalt enzyme. J. Biol. Chem. *249*, 4601–4607.

Iacobucci, G.A., Brose, D.J., Ray, R.J. and v. Eikeren, P. (1994) Enzymatic membrane method for the synthesis and separation of peptides. US-A 5,350,681.

Inouye, K. (1992) Effects of salts on thermolysin: Activation of hydrolysis and synthesis of N-carbobenzoxy-L-aspartyl-L-phenylalanine methyl ester, and a unique change in the absorption spectrum of thermolysin. J. Biochem. *112*, 335–340.

Inouye, K., Kuzuya, K. and Tonomura, B. (1998a) Sodium chloride enhances markedly the thermal stability of thermolysin as well as its catalytic activity. Biochim. Biophys. Acta *1388*, 209–214.

Inouye, K., Kuzuya, K. and Tonomura, B. (1998b) Effect of salts on the solubility of thermolysin: a remarkable increase in the solubility as well as the activity by the addition of salts without aggregation or dispersion of thermolysin. J. Biochem. *123*, 847–852.

Inouye, K., Lee, S. and Tonomura, B. (1996) Effect of amino acid residues at the cleavable site of substrates on the remarkable activation of thermolysin by salts. Biochem. J. *315*, 133–138.

Inouye, K., Minoda, M., Takita, T., Sakurama, H., Hashida, Y., Kusano, M. and Yasukawa, (2005) Extracellular production of recombinant thermolysin expressed in *Escherichia coli*, and its purification and enzymatic characterization. Protein Expr. Purif. *46*, 248–255.

Isowa, Y., Ohmori, M., Ichikawa, T., Mori, K., Nonaka, Y., Kihara, K., Oyama, K., Satoh, H. and Nishimura, S. (1979) The thermolysin-catalyzed condensation reactions of n-substituted aspartic and glutamic acids with phenylalanine alkyl esters. Tetrahedron Lett. *20*, 2611–2612.

Kamo, M., Inouye, K., Nagata, K. and Tanokura, M. (2005) Preliminary X-ray crystallographic analysis of thermolysin in the presence of 4 M NaCl. Acta Crystallogr. D Biol. Crystallogr. *61*, 710–712.

Kaplan, J. and DeGrado, W.F. (2004) De novo design of catalytic proteins. Proc. Natl. Acad. Sci. USA *101*, 11566–11570.

Kawamura, F. and Doi, R.H. (1984) Construction of a *Bacillus subtilis* double mutant deficient in extracellular alkaline and neutral proteases. J. Bacteriol. *160*, 442–444.

Khmelnitsky, Y.L., Budde, C., Arnold, J.M., Usyatinsky, A., Clark, D.S. and Dordick, J.S. (1997) Synthesis of water-soluble paclitaxel derivatives by enzymatic acylation. J. Am. Chem. Soc. *119*, 11554–11555.

Kidokoro, S. (1998) Design of protein function by physical perturbation method. Adv. Biophys. *35*, 121–143.

Kidokoro, S., Miki, Y., Endo, K., Wada, A., Nagao, H., Miyake, T., Aoyama, A., Yoneya, T., Kai, K. and Ooe, S. (1995) Remarkable activity enhancement of thermolysin mutants. FEBS Lett. *367*, 73–76.

Krix, G., Eichhorn, U., Jakubke, H.-D. and Kula, M.-R. (1997) Protease-catalyzed synthesis of new hydrophobic dipeptides containing non-proteinogenic amino acids. Enzyme Microb. Technol. *21*, 252–257.

Kubo, M. and Imanaka, T. (1988) Cloning and nucleotide sequence of the highly thermostable neutral protease gene from *Bacillus stearothermophilus*. J. Gen. Microbiol. *134*, 1883–1892.

Kubo, M., Mitsuda, Y., Takagi, M. and Imanaka, T. (1992) Alteration of specific activity and stability of thermostable neutral protease by site-directed mutagenesis. Appl. Environ. Microbiol. *58*, 3779–3783.

Kudryashova, E.V., Mozhaev, V.V. and Balny, C. (1998) Catalytic activity of thermolysin under extremes of pressure and temperature: modulation by metal ions. Biochim. Biophys. Acta *1386*, 199–210.

Kühn, D., Dürrschmidt, P., Mansfeld, J. and Ulbrich-Hofmann, R. (2002) Boilysin and thermolysin in dipeptide synthesis: a comparative study. Biotechnol. Appl. Biochem. *36*, 71–76.

Kunugi, S., Koyasu, A., Takahashi, S. and Oda, K. (1997) Peptide condensation activity of a neutral protease from *Vibrio* sp. T1800 (vimelysin). Biotechnol. Bioeng. *53*, 387–390.

Kuzuya, K. and Inouye, K. (2001) Effects of cobalt-substitution of the active zinc ion in thermolysin on its activity and active-site microenvironment. J. Biochem. *130*, 783–788.

Lee, K.H., Lee, P.M., Siaw, Y.S. and Morihara, K. (1992) Aspartame precursor synthesis in water miscible cosolvent catalysed by thermolysin. Biotechnol. Lett. *14*, 1159–1164.

Mansfeld, J., Petermann, E., Dürrschmidt, P. and Ulbrich-Hofmann, R. (2005) The propeptide is not required to produce catalytically active neutral protease from *Bacillus stearothermophilus*. Protein Expr. Purif. *39*, 219–228.

Mansfeld, J. and Ulbrich-Hofmann, R. (2000) Site-specific and random immobilization of thermolysin-like proteases reflected in the thermal inactivation kinetics. Biotechnol. Appl. Biochem. *32*, 189–195.

Mansfeld, J., Vriend, G., Dijkstra, B.W., Veltman, O.R., van den Burg, B., Venema, G., Ulbrich-Hofmann, R. and Eijsink, V.G.H. (1997) Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. J. Biol. Chem. *272*, 11152–11156.

Mansfeld, J., Vriend, G., van den Burg, B., Eijsink, V.G.H. and Ulbrich-Hofmann, R. (1999) Probing the unfolding region in a thermolysin-like protease by site-specific immobilization. Biochemistry *38*, 8240–8245.

Marie-Claire, C., Roques, B.P. and Beaumont, A. (1998) Intramolecular processing of prothermolysin. J. Biol. Chem. *273*, 5697–5701.

Marie-Claire, C., Ruffet, E., Beaumont, A. and Roques, B.P. (1999) The prosequence of thermolysin acts as an intramolecular chaperone when expressed in *trans* with the mature enzyme in *Escherichia coli*. J. Mol. Biol. *285*, 1911–1915.

Matsumiya, Y., Nishikawa, K., Aoshima, H., Inouye, K. and Kubo, M. (2004) Analysis of autodegradation sites of thermolysin and enhancement of its thermostability by modifying Leu^{155} at an autodegradation site. J. Biochem. *135*, 547–553.

Matthews, B.W. (1988) Structural basis of the action of thermolysin and related zinc peptidases. Acc. Chem. Res. *21*, 333–340.

Matthews, B.W., Jansonius, J.N., Colman, P.M., Schoenborn, B.P. and Dupourque, D. (1972) A structure for thermolysin. Nature *238*, 37–41.

Mejri, M., Pauthe, E., Larreta-Garde, V. and Mathlouthi, M. (1998) Effect of polyhydroxylic additives on the catalytic activity of thermolysin. Enzyme Microb. Technol. *23*, 392–396.

Miyanaga, M., Imamura, K., Sakiyama, T. and Nakanishi, K. (2000a) Kinetic analysis for synthesis of a dipeptide precursor using an immobilized enzyme in water-immiscible organic solvent. J. Biosci. Bioeng. *90*, 112–114.

Miyanaga, M., Ohmori, M., Imamura, K., Sakiyama, T. and Nakanishi, K. (2000b) Kinetics and equilibrium for thermolysin-catalyzed syntheses of dipeptide precursors in aqueous/organic biphasic systems. J. Biosci. Bioeng. *90*, 43–51.

Mock, W.L. and Aksamawati, M. (1994) Binding to thermolysin of phenolate-containing inhibitors necessitates a revised mechanism of catalysis. Biochem J. *302*, 57–68.

Mock, W.L. and Stanford, D.J. (1996) Arazoformyl dipeptide substrates for thermolysin. Confirmation of a reverse protonation catalytic mechanism. Biochemistry *35*, 7369–7377.

Moeschel, K., Nouaimi, M., Steinbrenner, C. and Bisswanger, H. (2003) Immobilization of thermolysin to polyamide nonwoven materials. Biotechnol. Bioeng. *82*, 190–199.

Murakami, Y. and Hirata, A. (1997) Continuous enzymatic synthesis of aspartame precursor at low pH using an extractive reaction. J. Ferment. Bioeng. *84*, 264–267.

Murakami, Y., Hirata, M. and Hirata, A. (1996) Mathematical approach to thermolysin-catalyzed synthesis of aspartame precursor. J. Ferment. Bioeng. *82*, 246–252.

Murakami, Y., Oda, T., Chiba, K. and Hirata, A. (2000b) Continuous production of N-(benzyloxycarbonyl)-L-glycyl-L-phenylalanine methyl ester utilizing extractive reaction in aqueous/organic biphasic medium. Prep. Biochem. Biotechnol. *30*, 15–22.

Murakami, Y., Yoshida, T. and Hirata, A. (1998) Enzymatic synthesis of N-formyl-L-aspartyl-L-phenylalanine methyl ester (aspartame precursor) utilizing an extractive reaction in aqueous/organic biphasic medium. Biotechnol. Lett. *20*, 767–769.

Murakami, Y., Yoshida, T., Hayashi, S. and Hirata, A. (2000a) Continuous enzymatic production of peptide precursor in aqueous/organic biphasic medium. Biotechnol. Bioeng. *69*, 57–65.

Nakanishi, K., Kamikubo, T. and Matsuno, R. (1985) Continuous synthesis of N(benzyloxycarbonyl)-L-aspartyl-L-phenylalanine methyl ester with immobilized thermolysin in an organic solvent. Bio/Technology *3*, 459–464.

Nakanishi, K., Takeuchi, A. and Matsuno, R. (1990) Long-term continuous synthesis of aspartame precursor in a column reactor with an immobilized thermolysin. Appl. Microbiol. Biotechnol. *32*, 633–636.

Oda, K., Takahashi, T., Takada, K., Tsunemi, M., Ng, K., Hiraga, K. and Harada, S. (2005) Exploring the subsite-structure of vimelysin and thermolysin using FRETS-libraries. FEBS Lett. *579*, 5013–5018.

O'Donohue, M.J. and Beaumont, A. (1996) The roles of the prosequence of thermolysin in enzyme inhibition and folding *in vitro*. J. Biol. Chem. *271*, 26477–26481.

Ooshima, H., Mori, H. and Harano, Y. (1985) Synthesis of aspartame precursor by solid thermolysin in organic solvent. Biotechnol. Lett. *7*, 789–792.

Park, Y.J., Liang, J.F., Ko, K.S., Kim, S.W. and Yang, V.C. (2003) Low molecular weight protamine as an efficient and nontoxic gene carrier: *in vitro* study. J. Gene Med. *5*, 700–711.

Park, C. and Marqusee, S. (2005) Pulse proteolysis: A simple method for quantitative determination of protein stability and ligand binding. Nat. Methods *2*, 207–212.

Patrick, W.M. and Firth, A.E. (2005) Strategies and computational tools for improving randomized protein libraries. Biomol. Eng. *22*, 105–112.

Pedersen, N.R., Halling, P.J., Pedersen, L.H., Wimmer, R., Matthiessen, R. and Veltman, O.R. (2002) Efficient transesterification of sucrose catalysed by the metalloprotease thermolysin in dimethylsulfoxide. FEBS Lett. *519*, 181–184.

Pelmenschikov, V., Blomberg, M.R. and Siegbahn, P.E. (2002) A theoretical study of the mechanism for peptide hydrolysis by thermolysin. J. Biol. Inorg. Chem. *7*, 284–298.

Persichetti, R.A., Clair, N.L.S., Griffith, J.P., Navia, M.A. and Margolin, A.L. (1995) Cross-linked enzyme crystals (CLECs) of thermolysin in the synthesis of peptides. J. Am. Chem. Soc. *117*, 2732–2737.

Raksakulthai, R., and Haard, N.F. (2003) Exopeptidases and their application to reduce bitterness in food: a review. Crtit. Rev. Food Sci Nutr. *43*, 401–445.

Rawlings, N.D., Tolle, D.P. and Barrett, A.J. (2004) MEROPS: the peptidase database. Nucleic Acids Res. 32 Database issue, D160–D164.

Rich, J.O., Mozhaev, V.V., Dordick, J.S., Clark, D.S. and Khmelnitsky, Y.L. (2002) Molecular imprinting of enzymes with water-insoluble ligands for nonaqueous biocatalysis. J. Am. Chem. Soc. *124*, 5254–5255.

Rival, S., Saulner, J. and Wallach, J. (2000) On the mechanism of action of pseudolysin: kinetic study of the enzymatic condensation of Z-Ala with Phe-NH_2. Biocatal. Biotrans. *17*, 417–429.

Roques, B.P., Noble, F., Dauge, V., Fournie-Zaluski, M.C. and Beaumont, A. (1993) Neutral endopeptidase 24.11: structure, inhibition, and experimental and clinical pharmacology. Pharmacol Rev. *45*, 87–146.

Saha, B.C. and Hayashi, K. (2001) Debittering of protein hydrolyzates. Biotechnol. Adv. *19*, 355–370.

Schechter, I. and Berger, A. (1967) On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. *27*, 157–162.

Schellenberger, A. and Ulbrich, R. (1989) Protein stabilization by blocking the native unfolding nucleus. Biomed. Biochim. Acta *48*, 63–67.

Stark, W., Pauptit, R.A., Wilson, K.S. and Jansonius, J.N. (1992) The structure of neutral protease from *Bacillus cereus* at 0.2-nm resolution. Eur. J. Biochem. *207*, 781–791.

Stöcker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Rüth, F.-X., McKay, D.B. and Bode, W. (1995) The metzincins – topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc peptidases. Protein Sci. *4*, 823–840.

Takagi, M., Imanaka, T. and Aiba, S. (1985) Nucleotide sequence and promoter region for the neutral protease gene from *Bacillus stearothermophilus*. J. Bacteriol. *163*, 824–831.

Thayer, M.M., Flaherty, K.M. and McKay, D.B. (1991) Three-dimensional structure of the elastase of *Pseudomonas aeruginosa* at 1.5-Å resolution. J. Biol. Chem. *266*, 2864–2871.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acid Res. *22*, 4673–4680.

Toma, S., Campagnoli, S., De Gregoriis, E., Gianna, R., Margarit, I., Zamai, M. and Grandi, G. (1989) Effect of Glu-143 and His-231 substitutions on the catalytic activity and secretion of *Bacillus subtilis* neutral protease. Protein Eng. *2*, 359–364.

Tran, L., Wu, X.C. and Wong, S.L. (1991) Cloning and expression of a novel protease gene encoding an extracellular neutral protease gene from *Bacillus subtilis*. J. Bacteriol. *173*, 6364–6372.

Ulbrich-Hofmann, R., Arnold, U. and Mansfeld, J. (1999) The concept of the unfolding region for approaching the mechanisms of enzyme stabilization. J. Mol. Catal. B: *7*, 125–131.

van den Burg, B., Eijsink, V.G.H., Stulp, B. and Venema, G. (1989) One-step affinity purification of *Bacillus* neutral proteases using Bacitracin-silica. J. Biochem. Biophys. Methods *18*, 209–219.

van den Burg, B., Eijsink, V.G.H., Stulp, B. and Venema, G. (1990) Identification of autodigestion target sites in *Bacillus subtilis* neutral proteinase. Biochem J. *272*, 93–97.

van den Burg, B., Enequist, H., van der Haar, M., Eijsink, V., Stulp, B. and Venema, G. (1991) A highly thermostable neutral protease from *Bacillus caldolyticus*: cloning and expression of the gene in *Bacillus subtilis* and characterization of the gene product. J. Bacteriol. *173*, 4107–4115.

van den Burg, B., Vriend, G., Veltman, O.R., Venema, G. and Eijsink, V.G.H. (1998) Engineering an enzyme to resist boiling. Proc. Natl. Acad. Sci. USA *95*,2056–2060.

van den Burg, B., de Kreij, A., van der Veek, P., Mansfeld, J. and Venema, G. (1999) Characterization of a novel stable biocatalyst obtained by protein engineering. Biotechnol. Appl. Biochem. *30*, 35–40.

van den Burg, B. and Eijsink,V. (2004) Thermolysin and related Bacillus metallopeptidases. In Handbook of Proteolytic Enzymes, 2nd ed. (Barrett, A.J., Rawlings, N.D. and Woessner, J.F. eds), 374–387, Elsevier, London.

Veltman, O.R., Eijsink, V.G.H., Vriend, G., de Kreij, A., Venema, G. and van den Burg, B. (1998) Probing catalytic hinge bending motions in thermolysin-like proteases by glycine→alanine mutations. Biochemistry *37*, 5305–5311.

Vercruysse, L., Smagghe, G., Herregods, G. and van Camp, J. (2005) ACE inhibitory activity on enzymatic hydrolysates of insect protein. J. Agric. Food Chem. *53*, 5207–5211.

Vieille, C. and Zeikus, G.J. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol. Mol. Biol. Rev. *65*, 1–43.

Vriend, G., Berendsen, H.J., van den Burg, B., Venema, G. and Eijsink, V.G.H. (1998) Early steps in the unfolding of thermolysin-like proteases. J. Biol. Chem. *273*, 35074–35077.

Vriend, G. and Eijsink, V.G.H. (1993) Prediction and analysis of structure, stability and unfolding of thermolysin-like proteases. J. Comput. Aided Mol. Des. *7*, 367–396.

Walsh, K.A., Burnstein, Y. and Pangburn, M.K. (1974) Thermolysin and other neutral metalloendopeptidases. Methods Enzymol. *34*, 435–440.

Wayne, S.I. and Fruton, J.S. (1983) Thermolysin-catalyzed peptide bond synthesis. Proc. Natl. Acad. Sci. USA *80*, 3241–3244.

Wetmore, D.R., Wong, S.L. and Roche, R.S. (1992) The role of the pro-sequence in the processing and secretion of the thermolysin-like neutral protease from *Bacillus cereus*. Mol. Microbiol. *6*, 1593–1604.

Ye, L., Ramström, O., Ansell, R.J., Mansson, M. and Mosbach, K. (1999) Use of molecularly imprinted polymers in a biotransformation process. Biotechnol. Bioeng. *64*, 650–655.

CHAPTER 15

AMINOPEPTIDASES

YOLANDA SANZ*

Departamento de Ciencia de los Alimentos, Instituto de Agroquímica y Tecnología de Alimentos, CSIC, Paterna, Valencia, Spain

**yolsanz@iata.csic.es*

1. INTRODUCTION

Aminopeptidases hydrolyse peptide bonds at the N-terminus of proteins and polypeptides whereas carboxypeptidases hydrolyse peptide bonds at the C-terminus. Omega peptidase is an additional term referring to special types of aminopeptidases and carboxypeptidases that are capable of removing terminal residues lacking a free α-amino or α-carboxyl group, or include linkages other than the α-peptide type (*e.g.* pyroglutamyl peptidases; McDonald and Barrett, 1986). Aminopeptidases can be subdivided into three groups: aminopeptidases in the strict sense which hydrolyse the first peptide bond in a polypeptide chain with the release of a single amino acid residue (aminoacyl- and iminoacyl peptidases [EC 3.4.11]); those that remove dipeptides or tripeptides (dipeptidyl- and tripeptidyl peptidases [EC 3.4.14]) from polypeptide chains; and those which only hydrolyse di- or tripeptides (dipeptidases [EC 3.4.15] and tripeptidases [EC 3.4.14.4]) (Sanderink *et al*, 1988). Aminopeptidases are widely distributed among bacteria, fungi, plants and mammals (Gonzales and Robert-Baudouy, 1996; Sanz *et al.*, 2002; Tu *et al*, 2003; Barret *et al*, 2004). Theses enzymes are located in different subcellular compartments including the cytoplasm, lysosomes and membranes, and can also be secreted into the extracellular medium. Based on catalytic mechanism, most of the aminopeptidases are metallo-enzymes but cysteine and serine peptidases are also included in this group. Though some aminopeptidases are monomeric, most show multimeric structures particularly those from eukaryotic organisms (McDonald and Barrett, 1986; Jones, 1991; Lowther and Matthews, 2002). The three-dimensional structures of some aminopeptidases have been solved, contributing to the understanding of their catalytic mechanism and

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 243–260

functions (Kim *et al.*, 1993; Bazan *et al.*, 1994; Joshua-Tor *et al.*, 1995; Lowther and Matthews, 2002). Aminopeptidases appear to act in concert with other peptidases to complete diverse proteolytic pathways. Thus, these enzymes can efficiently retrieve amino acids from dietary proteins and endogenous proteins degraded during protein turnover, thereby covering nutritional as well as other biological roles including protein maturation, hormone level regulation and cell-cycle control (McDonald and Barrett, 1986; Christensen *et al.*, 1999). Many of the mammalian enzymes play important functions in cellular processes involved in health and disease and, as a consequence, constitute targets for the pharmaceutical industry (Scornik and Botbol, 2001; Holz *et al.*, 2003; Rigolet *et al.*, 2005; Inguimbert *et al.*, 2005). Some aminopeptidases are also of great interest for their biotechnological and agro-industrial applications (Seppo *et al.*, 2003; FitzGerald and O'Cuinn, 2006).

2. CLASSIFICATION AND NOMENCLATURE OF AMINOPEPTIDASES

Aminopeptidases have been classified on the basis of their substrate specificity (broad or narrow), catalytic mechanism (metallo-, cysteine, and serine peptidases) and molecular structure (Gonzales and Robert-Baudouy, 1996; Barrett *et al.*, 2004; Rawling *et al.*, this volume). The nomenclature of many (aminoacyl or iminoacyl peptidases) has been determined by their preferences or requirements for a particular N-terminal amino acid. Thus, an enzyme that for instance showed its highest rate of hydrolysis on N-terminal methionyl bonds was named methionyl aminopeptidase or aminopeptidase M. In an attempt to avoid ambiguity, the subcellular location (membrane, microsomal or cytosolic) has also been used to name aminopeptidases having similar specificities. In the cases of di- and tripeptidases their names have been based on substrate size requirements. In addition, the names of dipeptidyl (DPP) and tripeptidyl peptidases (TPP) were followed by a Roman number to differentiate between the various types described and this numbering convention has been retained. In the nomenclature of peptidases identified in lactic acid bacteria the term 'Pep' is used when the corresponding peptidase gene sequence is known, followed by a capital letter indicating the specificity and homology to other known peptidases, *e.g.* PepN for the homologue of aminopeptidase N (Tan *et al.*, 1993). Nevertheless, alternative names and abbreviations often appear in the literature. Recently, the MEROPS peptidase information database (http://www.merops.sanger.ac.uk; see chapter by Rawlings *et al.* in this volume) created a hierarchical structure-based classification of peptidases into families and clans. Members of a family are homologues, and families that are thought to be homologous are grouped together into clans. Clans consist of families of peptidases that share a single evolutionary origin, evidenced by similarities in their tertiary structures and/or the order of catalytic-site residues and common sequence motifs around the catalytic residues.

3. ANIMOPEPTIDASE TYPES AND FUNCTIONS

3.1. Mammalian Aminopeptidases

Aminopeptidases were among the first proteases to be discovered in mammalian tissues and a large number have already been characterized (Barret *et al*., 2004). The best characterized mammalian aminopeptidases and their biochemical properties are shown in Table 1. On the basis of specificity they can be divided into different groups: (i) aminopeptidases of broad specificity (*e.g.* leucyl aminopeptidase, and membrane and cytosol alanyl aminopeptidases); (ii) aminopeptidases of narrow specificity with preference for basic amino acid residues (aminopeptidase B), acid amino acid residues (glutamyl and aspartyl aminopeptidases), cysteine (cystinyl aminopeptidase), methionine (methionyl aminopeptidase), or bonds containing proline (prolyl aminopeptidase and aminopeptidase P); and (iii) dipeptidyl peptidases (DPP I, II, III, and IV) and tripeptidyl peptidases (TPP I and II) that release di- and tripeptides; respectively (Cunningham *et al*., 1997; Sanz *et al*., 2002; Albiston

Table 1. Main types and properties of mammalian aminopeptidases

Enzyme	EC number	Catalytic type	Specificity	Family
Aminopeptidase (AP)			$X- \Downarrow -Y-(Z)_n$	
Leucyl AP	3.4.11.1	Metallo	X = Leu	M17
Membrane alanyl AP	3.4.11.2	Metallo	X = Ala, Phe, Tyr, Leu	M1
Cytosol alanyl AP	3.4.11.14	Metallo	X = Ala	M1
Aminopeptidase B	3.4.11.6	Metallo	X = Arg, Lys	M1
Glutamyl AP	3.4.11.7	Metallo	X = Glu, Asp	M1
Cystinyl AP	3.4.11.3	Metallo	X = Cys	M1
Methionyl AP	3.4.11.18	Metallo	X = Met	M24A
Aminopeptidase P	3.4.11.9	Metallo	X and Y = Pro	M24B
Prolyl AP (PIP)	3.4.11.5	Serine	X = Pro	S33
Bleomycin hydrolase	3.4.22.40	Cysteine	X = Met, Leu, Ala Bleomycin peptide	C1B
Dipeptidyl peptidases (DPP)			X-Y-⇓-(Z)n	
DPP I	3.4.14.1	Cysteine	X ≠ Arg or Lys, Y or Z ≠ Pro	C1
DPP II	3.4.14.2	Serine	Y = Ala or Pro	S28
DPP III	3.4.14.4	Metallo	X-Y = Arg-Arg	M49
DPP IV	3.4.14.5	Serine	Y = Pro	S9B
Tripeptidyl peptidases (TPP)			X-Y-T-⇓-(Z)n	
TPP I	3.4.14. 9	Serine	Gly-Pro-T = hydrophobic	S53
TPP II	3.4.14.10	Serine	Ala-Ala-Phe T or Z ≠ Pro	S8

et al., 2004). The characteristics of the main aminopeptidases of broad and narrow specificities and proline-specific peptidases are briefly reviewed.

3.1.1. Mammalian aminopeptidases of broad specificity

Leucyl aminopeptidase (LAP) is a ubiquitous enzyme that has also been referred to as cytosol aminopeptidase and leucine aminopeptidase. It was the first cytosolic aminopeptidase to be identified (Linderstrom-Lang, 1929). LAP preferentially releases Leu located as the N-terminal residue of peptides, and can also release other amino acids including Pro but not Arg or Lys. This enzyme is a hexamer of identical chains and has a molecular mass of 324-360 (Kohno *et al*, 1986). Human LAP is involved in the breakdown of the peptide products of intracellular proteinases and is one of the enzymes that trims proteasome-produced peptides for presentation by the major histocompatibility complex class I molecules. Expression of the encoding gene is promoted by interferon gamma (Beninga *et al*, 1998).

Membrane alanyl aminopeptidase has also been referred to as aminopeptidase M due to its membrane localization since there is a cytosolic counterpart, and also as aminopeptidase N due to its preference for neutral amino acids. The amino acid residue preferentially released is Ala, but most amino acids including Pro may be hydrolysed by this enzyme. When a terminal hydrophobic residue is followed by Pro, the two may be released as an intact X-Pro dipeptide (McDonald and Barret, 1986). In most species the native enzyme is a homodimer with a molecular mass of 280-300 and is glycosylated. The mammalian enzyme plays a role in the final digestion of peptides generated from proteins by gastric and pancreatic protease hydrolysis. It is also important for the inactivation in the kidney of blood-borne peptides such as enkephalins and the neuropeptide ‘substance P’. Furthermore, it is a regulator of IL-8 bioavailability in the endometrium and therefore may contribute to the regulation of angiogenesis. This aminopeptidase is also the myeloid leukaemia marker CD13 and serves as a receptor for human coronavirus (Shimizu *et al*, 2002; Albiston *et al*, 2004).

3.1.2. Mammalian aminopeptidases of narrow specificity

Glutamyl aminopeptidase is also referred to as aminopeptidase A, angiotensinase A and aspartate aminopeptidase. It releases N-terminal Glu (and to a lesser extent Asp) from a peptide. It is generally a membrane-bound enzyme involved in the formation of the brain heptapeptide angiotensin III which exerts a tonic stimulatory effect on the central control of blood pressure and is a regulator of blood vessel formation (Fournie-Zaluski *et al*, 2004.).

Methionyl aminopeptidase has also been named methionine aminopeptidase and peptidase M. It releases N-terminal amino acids, preferentially methionine, from peptides but only when the second residue is small and uncharged. In eukaryotes, two types of methionyl aminopeptidases exist due to protein synthesis occurring in the mitochondria (type I) and in the cytoplasm (type II). The type I peptidase is similar to the bacterial methionyl aminopeptidase whereas type II resembles the enzyme from archaea (Arfin *et al*, 1995). The mammalian enzymes are involved in

the regulation of protein synthesis and in the processing of those proteins required for the formation of new blood vessels in normal development, tumour growth and metastasis (Yeh *et al.*, 2000; Selvakumar *et al.*, 2005; Zhong *et al.*, 2006).

3.1.3. Proline-specific peptidases

This group includes two aminopeptidases: prolyl aminopeptidase and X-Pro aminopeptidase. The prolyl aminopeptidase has variously been named Pro-X aminopeptidase, proline aminopeptidase and proline iminopeptidase. It releases N-terminal proline from a peptide. This enzyme was first detected in *E. coli* but is widely distributed in nature and also present in the cytosol of mammalian cells (Matsushima *et al.*, 1991). In contrast to the bacterial form, the mammalian enzyme is not specific for prolyl bonds (Cunningham *et al.*, 1997). X-Pro aminopeptidase has also been termed aminopeptidase P and proline aminopeptidase. It releases any N-terminal amino acid residue, including proline, from oligopeptides and even dipeptide and tripeptides in which the penultimate N-terminal residue is proline. The preferred substrates have a hydrophobic or basic residue at the N-terminus. The mammalian enzyme exists in membrane-bound and cytosolic forms (Cottrell *et al.*, 2000). It appears to contribute to the processing of bioactive peptides involved in the cardiovascular and pulmonary systems, and the degradation of collagen products (Yaron and Naider, 1993; Yoshimoto *et al.*, 1994).

Of the mammalian peptidyl peptidases, dipeptidyl peptidase IV (DPP IV) is the best know. It releases an N-terminal dipeptide from polypeptides in which, the penultimate residue is Pro (preferentially but not exclusively) and provided that the antepenultimate residue is neither Pro nor hydroxyproline (Leiting *et al.*, 2003). This enzyme is anchored to the cell membrane and expressed in various cell types. It has a calculated molecular mass of 88 kDa and the native enzyme is a homodimer. DPP IV plays a key role in various regulatory processes, acting on a number of bioactive oligopeptides including neuropeptides, endomorphins, circulating peptide hormones, glucagon-like peptides (GLP-1 and GLP-2), gastric inhibitory peptide (GIP) and paracrine chemokines, leading to modification of their biological activities or even their inactivation (Augustyns *et al.*, 2005).

3.2. Microbial Aminopeptidases

The first studies on microbial aminopeptidases were carried out over 40 years ago, and since then a large number of aminopeptidases of microbial origin have been characterized (Gonzales and Robert-Baudouy, 1996; Jones, 1991; Kunji *et al.*, 1996; Christensen *et al.*, 1999; Sanz and Toldrá, 2002; Sanz *et al.*, 2002; Barret *et al.*, 2004; Nampoothiri *et al.*, 2005; Savijoki *et al.*, 2006). The main types of microbial activity characterized to date as well as their properties are summarized in Table 2. These enzymes can be divided according to their specificities into groups similar to those described for the mammalian enzymes: (i) general aminopeptidases showing broad specificity (PepN, PepC, LAP and PepS); (ii) aminopeptidases of narrow specificity that selectively hydrolyse certain amino acid residues such

Table 2. Main types and properties of microbial aminopeptidases

Enzyme	EC number/ homologous	Catalytic type	Specificity	Family
Aminopeptidases (AP)			X- ⇓-(Y)-$(Z)_n$	
PepN/ Lysyl aminopeptidase	Mammalian AP N	Metallo	X = Lys, Leu,	-
Bleomycin hydrolase GAL6/BLH1/YCP1	3.4.22.40	Cysteine	X = Arg, Tyr Bleomycin	C1B
PepC/aminopeptidase C	Bleomycin hydrolase	Cysteine	X = Lys, Glu, Ala, Met, Leu X or Y≠ Pro	C1B
PepA/aminopeptidase A	-	Metallo	X = Glu, Asp, Ser	M42
PepS	-	Metallo	X = Arg, Trp	M29
CAP/PepA Leucine aminopeptidase	3.4.11.10	Metallo	X = Leu, Met, Phe, Arg	M17
Aminopeptidase Y/yscl	3.4.11.15	Metallo	X = Arg, Lys, Ala	M28
Aminopeptidase M/MAP	3.4.11.18	Metallo	X = Met	M24A
PepP/AminopeptidaseP	3.4.11.9	Metallo	Y = Pro	M24B
PepI/Proline iminopeptidase	3.4.11.5	Serine	X = Pro	S33
D-stereospecific aminopeptidase/DppA	-	Serine	X = D-Ala, D-Ser or D-Thr	S12
Dipeptidases			X-⇓-Y	
PepV/peptidase V	-	Metallo	X = Lys, Leu, Met βAla-dipeptides	M20A
PepD/PepDA	-	Cystein	X = Lys, Met, Leu	C69
PepQ/prolidase	3.4.13.19	Metallo	Y = Pro	M24B
PepR/prolinase	-	Serine	X = Pro	-
Tripeptidases			X-⇓-Y-Z	
PepT/Peptidase T	-	Metallo	Leu-Gly-Gly	-
Dipeptidyl-peptidases			X-Y-⇓-$(Z)_n$	
PepX/X-Pro-dipeptidyl peptidase	3.4.14.11	Serine	Y = Pro	S15

as acidic residues (PepA) and methionine (MAP), D-amino acid residues (DppA) or peptide bounds containing proline (PepI and PepP); (iii) dipeptidases hydrolysing peptide bounds containing proline (PepQ and PepR); (iv) dipeptidases (PepV and PepDA) and tripeptidases (PepT) of broad specificity that only hydrolyse dipeptides or tripeptides, respectively; and (v) dipeptidyl peptidases showing specificity for N-terminal X-Pro.

Microbial aminopeptidases play important roles in the utilization of exogenous proteins as a source of essential amino acids that can be utilized for protein synthesis, the generation of metabolic energy and the recycling of reduced cofactors (Christensen *et al*., 1999). They are also implicated in the final steps of protein turnover and in more specific cellular functions such as the processing of newly synthesized proteins and high copy number plasmid stabilization (Gonzales and Robert-Baudouy, 1996).

3.2.1. *Microbial aminopeptidases of broad specificity*

PepN aminopeptidases, also known as lysyl aminopeptidases, have been identified in numerous bacterial species (*e.g. E. coli, Pseudomonas*, and lactic acid bacteria. Gonzales and Robert-Baudouy, 1996; Christensen *et al*, 1999). In most micro-organisms these are monomeric enzymes of about 95 kDa. Their primary sequences are homologous to mammalian aminopeptidase N and conserve the signature sequence of zinc-dependent metallo-peptidases (Gonzales and Robert-Baudouy, 1996; Kunji *et al*, 1996). In lactic acid bacteria this enzyme is involved in the utilization of caseins as an exogenous source of amino acids (Kunji *et al*, 1996).

LAPs (PepA in *E. coli*) are also zinc-metallo aminopeptidases of broad specificity identified in Gram-negative bacteria and fungi (Gonzales and Robert-Baudouy, 1996; Nampoothiri *et al*, 2005). These enzymes show sequence homology with bovine lens leucine aminopeptidase and similar specificity (Gonzales and Robert-Baudouy, 1996). Pep L aminopeptidases identified and partially characterized in different species of *Lactobacillus* seem, however, to be serine peptidases (Sanz *et al*, 1997; Christensen *et al*, 1999). PepC and bleomycin hydrolases are cysteine aminopeptidases of relatively broad specificity identified in lactic acid bacteria and yeast, respectively. Both exhibit similarity to mammalian bleomycin hydrolases (Kunji *et al*, 1996).

3.2.2. *Aminopeptidases of narrow specificity*

MAPs of microbial origin show high levels of similarity with mammalian MAP, conserving all five metal-binding residues and also maintaining similar specificity. The enzymes from prokaryotes and yeasts seem to be monomers of 29 kDa and 44 kDa, respectively. They play critical biological roles since their inactivation in *E. coli, S. typhimurium* and *S. cerevisiae* result in lethal phenotypes (Gonzales and Robert-Baudouy, 1996). A homologue to mammalian MAP type 2 is also present in yeast and shows subtle differences in its peptide substrate specificity (Chen *et al.*, 2002).

Aminopeptidase A, also referred to as glutamyl aminopeptidase and PepA, was identified in *Lactococcus lactis*. The genetic and physicochemical properties of this enzyme are not related to other aminopeptidases in prokaryotes or eukaryotes of similar specificity except for the enzyme purified from *Streptococcus thermophilus*. It specifically hydrolyses Glu and Asp residues, and to a lesser extent Ser residues, from the N-terminus of oligopeptides. In most cases the native enzyme seems to be a hexamer with a molecular mass of 240 kDa although other values (440–520) have been reported. The lactococcal enzyme was not demonstrated to be essential for growing on milk caseins but is thought to be important for flavour generation in dairy products (l'Anson *et al*, 1995).

3.2.3. *Proline-specific aminopeptidases*

A set of peptidases specialised in the hydrolysis of proline-containing peptides has been detected in lactic acid bacteria, and is thought to be necessary for the complete degradation of caseins since they have a high proline content (Kunji *et al*, 1996).

This group includes: two aminopeptidases (PepP and PepI), two dipeptidases (PepQ and PepR) and a dipeptidyl-peptidase (PepX), all of which conserve the consensus signatures of their catalytic types (metallo or serine peptidases). PepX has also been detected in streptococci and is thought to play a role in the pathological processes caused by *Streptococcus gordonii* and *S. agalactiae*, such as endocarditis, neonatal sepsis and meningitis (Rigolet *et al*, 2005).

3.3. Plant Aminopeptidases

Several aminopeptidases have also been identified in plants. These enzymes are believed to play biological roles in protein turnover, stress responses, protein mobilization from cotyledons after germination, protein maturation and meiosis. The aminopeptidases identified in plants also include enzymes of the broad and narrow specificities previously described. Among them, the LAP from tomato (*Lycopersicon esculentum*) is one of the best characterized animopeptidases of broad specificity. At least two distinct LAPs have been identified which seem to have different expression patterns and roles. The best-known enzyme (LAPa-A) is a wound-induced metallo aminopeptidase which preferentially hydrolyses substrates having N-terminal Leu, Arg or Met residues and with a homo-hexamer structure (Tu *et al*, 2003). Aminopeptidase N has also been identified in cucumber (*Cucumis sativus L. suyo*) and *Arabidopsis thaliana*. This is a metalloenzyme with similar specificity and sequence homology to that of aminopeptidases N which is classified into family M1 (Yamauchi *et al*, 2001). Among the aminopeptidases of narrow specificity, methionine aminopeptidases have also been identified in diverse plant species such as *Arabidopsis thaliana,* and are required for normal plant development (Ross *et al*, 2005). Aminopeptidase P has been identified in tomato (*Lycopersicon esculentum*) and is more than 40% identical to mammalian aminopeptidase P. It hydrolyses the amino terminal X-Pro bonds of bradykinin and also shows some endoproteolytic activity (Hauser *et al*, 2001).

4. CATALYTIC MECHANISM AND STRUCTURE OF AMINOPEPTIDASES

4.1. Metalloaminopeptidases

Metalloaminopeptidases, which constitute the largest group of aminopeptidases, are hydrolases in which the nucleophilic attack on a peptide bond is mediated by a water molecule that is activated by a divalent metal cation (Barret *et al*, 2004). Some aminopeptidases require a single metal ion for catalysis (*e.g.* MAP) while others require two metal ions (*e.g.* LAP from bovine lens; Lowther and Matthews, 2002). The known metal ligands of metallopeptidases are His, Glu, Asp or Lys residues. In addition to these metal ligands at least one additional residue, which can be Glu, Lys or Arg, is required for catalysis (Barret *et al*, 2004). Despite the differences in structure and metal centres among the metalloaminopeptidases, overall they utilize a

similar reaction mechanism. The carbonyl group of the substrate binds to the active site interacting with metal site 1 and a conserved enzyme residue. The N-terminus of the substrate also interacts either with metal site 2 or with one or more acidic enzyme residues. The scissile peptide bond is attacked by a solvent molecule that has been activated by its interaction with the metal ion and an enzyme residue that functions as a general base. Breakdown of the intermediate is most likely promoted by the addition of a proton to the leaving amino group donated by the general base. Differences in the binding pockets are responsible for the differences in substrate specificity, being broad or restrictive. Conserved amino acid side chains and the backbone atoms that are adjacent to the metal centre also provide key interactions. The oligomeric nature of some of the active enzymes also appears to be important for substrate specificity (Lowther and Matthews, 2002; Holz *et al*, 2003).

Metalloaminopeptidases have been subdivided into six clans (MA, MF, MG, MH, MN and MQ) on the basis of their folds, their active site architectures and the identities of active metal ions (see Rawling *et al.* in this volume). The best-known metalloaminopeptidases are found in clans MA (subclan MA(E)) – those enzymes which have only one catalytic metal ion -, and MF and MG, which have co-catalytic metal ions.

Amongst others, subclan MA(E) contains zinc-dependent peptidases which belong to peptidase family M1 and include bacterial lysyl aminopeptidase (PepN), and the mammalian enzymes membrane alanyl aminopeptidase (aminopeptidase N) and leukotriene A4 hydrolase, the latter possessing aminopeptidase and epoxyhydrolase activities. The peptidases of family M1 have a conserved His-Glu-X-X-His (HEXXH) motif involved in catalysis; they are also dependent on a single zinc ion for activity. The catalytic zinc ion is bound by the two histidines in the motif and the glutamate is a catalytic residue. The tertiary structures of members of this family show a two-domain structure with the active site in the cleft between them (Turner *et al.*, 2004). The structure of leukotriene A4 hydrolase has been solved revealing a three-domain protein in which the catalytic domain is the middle one. This domain contains an antiparallel β-sheet and α-helices, similar to that of thermolysin which is the type example of subclan MA(E) (Thunnissen *et al*, 2001).

Clan MF is comprised of peptidase family M17 that includes LAPs from eukaryotes and bacteria (PepA). These enzymes require co-catalytic metal ions for activity (Barret *et al*, 2004). The three-dimensional structure of bovine lens LAP has been solved (Burley *et al*, 1990; Kim *et al*, 1993; Cappiello *et al*, 2006), revealing that the protein is a homohexamer and that each monomer contains two domains: the N-terminal and the catalytic C-terminal domain, the latter containing the metal centre. Both domains contain α and β structures, with β-sheets in an α/β/α layering. The monomers within the hexamer are arranged as two layers of trimers. The two metal ions Zn1 and Zn2 are coordinated by the side chains of conserved amino acid residues of the enzyme (Lipscomb and Sträter, 1996). Zn2 binds the N-terminus of the substrate; Zn1 is also thought to provide critical binding and stabilizing interactions for the substrate and transition stages (Sträter and Lipscomb, 1995). The three-dimensional structure of the *E. coli* enzyme (Fig. 1)

has also been solved showing a hexameric quaternary structure similar to that of bovine lens LAP, but containing two manganese ions in the active site (Sträter *et al.*, 1999).

Clan MG includes peptidases of family M24 which is itself split into two subfamilies: M24A which includes the methionyl aminopeptidases and M24B which includes the aminopeptidase P and prolidase (X-Pro dipeptidase or PepQ) type peptidases. They have two cobalt or two manganese ions in their active centres. The narrow specificity of these enzymes is related with a common pitta-bread fold which contains a metal centre flanked by well-defined substrate binding pockets (Bazan *et al.*, 1994). The structure of the *E. coli* enzyme revealed the two metal ions to be sandwiched between two β-sheets surrounded by four α helices, yielding a structure with pseudo-2-fold symmetry (Roderick and Matthews, 1993). The restricted specificity suggests that these enzymes play roles in regulatory processes rather than in general protein degradation (Lowther and Matthews, 2002).

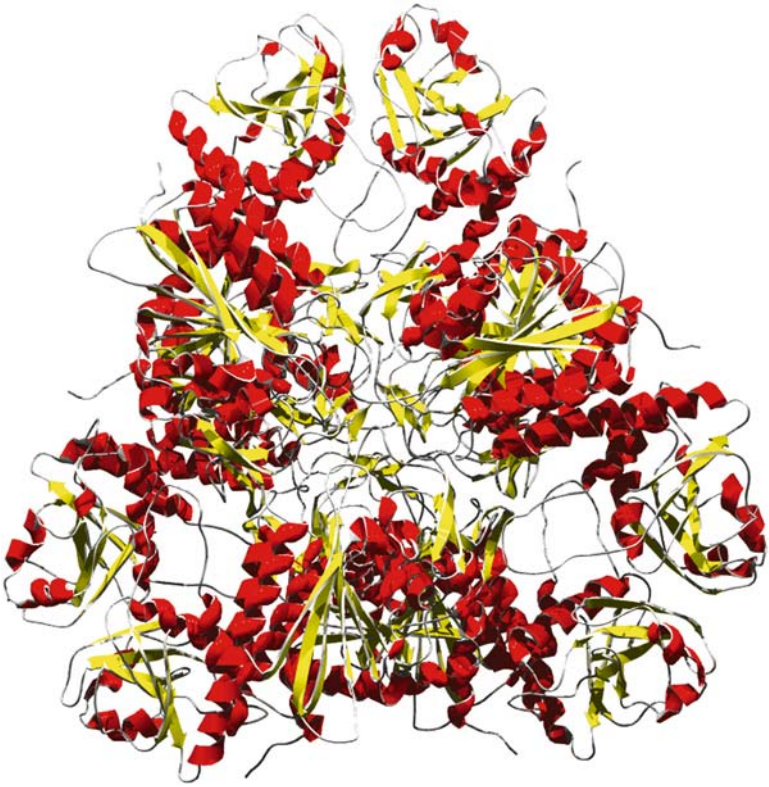


Figure 1. Overall structure of hexameric *E. coli* leucyl aminopeptidase (Sträter *et al.*, 1999)

4.2. Cysteine and Serine Aminopeptidases

Cysteine and serine aminopeptidases have no ionic co-factors associated with their structures. Catalysis requires a highly reactive cysteine or serine residue. In both cases the reaction begins with a nucleophilic attack on the carbon of the carbonyl group involved in the peptide bond of the substrate. In the case of cysteine aminopeptidases the attack is made by the sulphur of the sulphydryl group whereas in the serine aminopeptidases the attack is made by the oxygen of the hydroxyl group (Gonzales and Robert-Baudouy, 1996).These types of enzymes are less abundant than the metalloaminopeptidases and include cysteine peptidases of relatively broad specificity such as bleomycin hydrolase and PepC, and serine peptidases of narrow specificity such as proline-specific peptidases (PepI or prolyl aminopeptidase, PepX and DPP IV).

Cysteine aminopeptidases are included in clan CA and family C1B. They show the signature sequences of the catalytic site of the papain superfamily, and the amino acid residues important for catalysis (Gln, Cys, His, and Asn/Asp). The crystal structures of yeast bleomycin hydrolase (GAL6) and the human enzyme have been solved and show overall similarity (Zheng *et al*., 1998; Joshua-Tor *et al.*, 1995; O'Farrell *et al*., 1999). The proteins are hexameric, the six identical subunits forming barrel structures with the active sites embedded in a prominent central channel (Zheng *et al*., 1998). The monomers have a papain-like polypeptide fold as the core, with additional structural and functional modules inserted into loop regions. The crystallographic model of *Lactococcus lactis* PepC reveals that it is a homohexamer the subunits of which leave a narrow channel restricting the access to peptides. The projection of the C-terminal arm into the active site is a major difference relative to papain which, together with the overall architecture of the hexamer, limits the access to the active site cleft and may explain why peptidase activity observed *in vitro* has been restricted to small peptides (Joshua-Tor *et al.*, 1995). This carboxyl-terminal arm, also conserved in bleomycin hydrolases, is critical for oligomerization and aminopeptidase activity but not for endopeptidase activity (Mistou *et al*., 1994; Joshua-Tor *et al*., 1995; Mata *et al*., 1999).

Serine aminopeptidases do not belong to the main group of serine proteolytic enzyme families represented by trypsin and subtilisin. Peptide sequence analysis revealed that both prolyl aminopeptidase (PIP), PepI, DPP IV and PepX contain a catalytic triad which consists of Ser, His and Asp, and are related to prolyl oligopeptidases (Engel *et al*., 2005). The three-dimensional structures of the PIPs of several bacteria have been solved (Yoshimoto *et al*., 1999; Engel *et al*., 2005). The PIP protein is folded into two contiguous domains. The larger domain shows the general topology of the α/β hydrolase fold, with a central eight-stranded β-sheet flanked by two helices and the 11 N-terminal residues on one side, and by four helices on the other. The smaller domain is located above the larger and consists of six helices (Fig. 2). The catalytic triad (Ser 113, His 296, and Asp 268) is located near the large cavity at the interface between the two domains. The residues which make up the hydrophobic pocket line the smaller domain, and the specificity of the exo-type enzyme originates from this smaller domain (Yoshimoto *et al*., 1999).

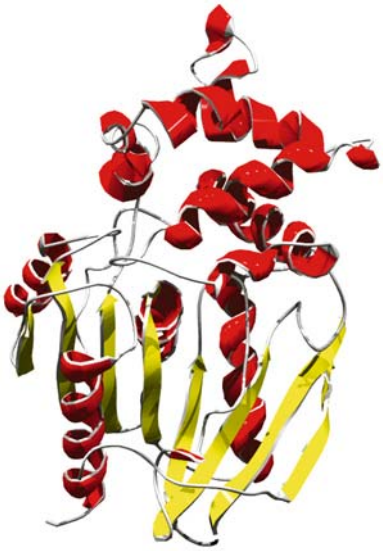


Figure 2. Monomer of prolyl aminopeptidase from *Serratia marcescens* showing its two distinct domains: the larger α/β domain in the bottom part of the figure and the smaller one, composed of six helices, on top (Yoshimoto *et al*, 1999)

The crystal structures of lactococcal PepX and a number of mammalian DPP IV enzymes have also been solved as a result of the interest generated by their key roles in diverse regulatory processes and the therapeutic potential of DPP IV inhibitors (Engel *et al*, 2003). The mammalian enzyme is an α/β-hydrolase that is secreted as a mature monomer but requires oligomerization to display normal proteolytic activity. Each monomer (Fig. 3) consists of an N-terminal β-propeller domain and an α/β-hydrolase domain enclosing an internal cavity that harbours the active site. The cavity is connected with the external environment through two different openings, the "propeller opening" and a "side opening" (Engel *et al*, 2005).The lactococcal enzyme is a homodimer with 2-fold symmetry. It folds into four distinct contiguous domains. The α/β-hydrolase fold is the largest domain and contains the catalytic site. The shortest domain is involved in oligomerization and binding specificity (Chich *et al*, 1995).

5. INDUSTRIAL APPLICATIONS OF AMINOPEPTIDASES

5.1. The Pharmaceutical Industry

Aminopeptidases play important roles in diverse cellular processes. As a consequence, pharmaceutical applications are being directed to control their activity in pathophysiological processes as well as the development of diagnosis tools and markers of physiological pathways (Brown, 2005). Most of the applications

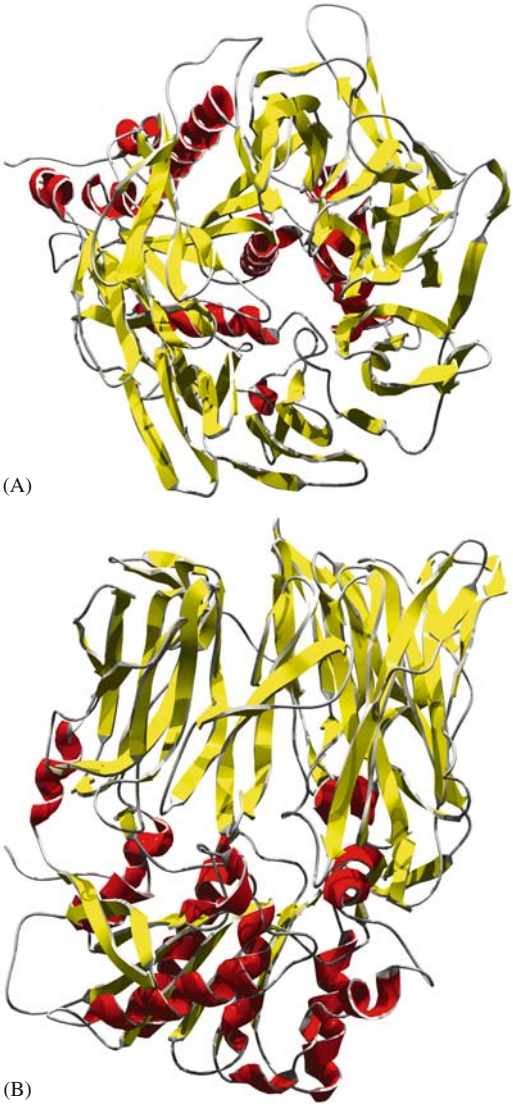


Figure 3. Mammalian (pig kidney) dipeptidyl peptidase (Engel *et al*., 2003). Panel **A** shows a view of the protein facing the N-terminal β-propeller domain. Panel **B** represents the protein from a perpendicular orientation showing the β-propeller domain in the upper part and α/β domain harbouring the catalytic site at the bottom

are oriented to the design of inhibitors for specific aminopeptidases. Selective inhibitors of glutamyl aminopeptidase (aminopeptidase A) constitute potential anti-hypertensive agents due to the role of this enzyme in the conversion of angiotensin II into angiotensin III, which plays an essential role in control of arterial blood

pressure (Cogolludo *et al*., 2005; Inguimbert *et al*., 2005). The design of inhibitors of methionine aminopeptidases is also considered to be of therapeutic potential due to the role of these enzymes in angiogenesis and tumour growth (Selvakumar *et al.*, 2005; Zhong and Bowen, 2006). Inhibitors of the expression of alanyl aminopeptidase (aminopeptidase N), which is deregulated in inflammatory diseases, cancer, leukaemia, diabetic nephropathy and rheumatoid arthritis, are also being developed to try to control these disorders (Bauvois and Dauzonne, 2006; Ansorge *et al.*, 2006). The design of inhibitors of DPP IV and related proline-specific peptidases is currently under investigation since these enzymes are involved in peptide metabolism of members of the PACAP/glucagon peptide family, neuropeptides and chemokines. The most promising applications of these agents are in the treatment of type 2 diabetes and immunological disorders (Augustyns *et al*., 2005; Mest, 2006). The inhibition of other aminopeptidases such as PepX (involved in infections by *Streptococcus gordonii*), the stereospecific DppA aminopeptidase (involved in peptidoglycan synthesis) and methionyl aminopeptidase, also constitute potential pharmaceutical targets to control microbial infections (Holz *et al*., 2003; Rigolet *et al.*, 2005; Schiffmann *et al*., 2006).

5.2. Biotechnological and Food Industrial Applications

One of the main industrial applications of aminopeptidases and their microbial producer strains is the manufacture of protein hydrolysates and protein-rich fermented products derived from soy, meat, milk, cereals, *etc*. (Meyer-Barton *et al.*, 1994; Suchiibun *et al*., 1993; Chevalet *et al*., 2001; Scharf *et al*., 2006). Food protein hydrolysates are manufactured for diverse purposes such as the fortification of foods and beverages, the elaboration of pre-digested ingredients for enteral/parenteral nutrition, and the generation of bioactive peptides and healthcare products (FitzGerald and O'Cuinn, 2006). The use of animopeptidases in these industrial processes not only contributes to the improvement of nutritional value but also the flavour of the final product by promoting the degradation of hydrophobic peptides which have undesirable tastes and the release of other peptides of agreeable taste characteristics and free amino acids. The application of these strategies to cheese ripening has been thoroughly investigated due to the high content of hydrophobic amino acid residues (*e.g.* proline) present in milk caseins (Meyer-Barton *et al.*, 1994; Savijoki *et al*., 2006). The use of proline-specific peptidases together with aminopeptidases of broad specificity (*e.g.* LAP) has been especially successful in the food industry (Raksakulthai and Haard, 2003). Some of the commercial aminopeptidases that are used to reduce bitterness in food are LAPs from lactic acid bacteria, *Rhizopus oryzae, Aspergillus oryze* and *Aspergillus sojae* (Nampoothiri *et al*., 2005). The use of lactic acid bacteria expressing specific peptidase activities during food protein processing is also being explored for reducing the levels of toxic and allergenic epitopes present in milk and cereal proteins (Di Cagno *et al*., 2004). A similar approach has also been used for the generation of bioactive peptides with antihypertensive, immunomodulatory and

antimicrobial properties (Meisel, 2004). Recently, the peptidases of *Lactobacillus helveticus* R211 and R389 have been found to generate casein-derived peptides that inhibit the angiotensin converting enzyme and are active *in vivo* (Leclerc *et al.*, 2001; Seppo *et al.*, 2003).

The application of combinations of peptidases to hydrolyse collagen for cosmetic uses has also been developed (Shigeri *et al.*, 2005). In addition, thermostable high-activity aminopeptidases constitute alternatives for biotechnological applications such as the processing of recombinant proteins (Gilboa *et al.*, 2001).

REFERENCES

Albiston, A.L., Ye, S. and Chai, S.Y. (2004) Membrane bound members of the M1 family: more than aminopeptidases. Protein Pept. Lett. *11*, 491–500.

Ansorge, S., Bank, U., Nordhoff, K., Taeger, M. and Striggow, F. (2006) Dual alanyl aminopeptidase and dipeptidyl peptidase IV inhibitors for functionally influencing different cells and for treating immunological, inflammatory, neuronal and other diseases. Patent Application WO2005034940.

Arfin, S.M., Kendall, R.L., Hal, L., Weaver, L.H., Stewart, A.E., Matthews, B.W. and Bradshaw, R.A. (1995) Eukaryotic methionyl aminopeptidases: two classes of cobalt-dependent enzymes. Proc. Natl. Acad. Sci. U S A. *92*, 7714–7718.

Augustyns, K., Van der Veken, P., Senten, K. and Haemers, A. (2005) The therapeutic potential of inhibitors of dipeptidyl peptidase IV (DPP IV) and related proline-specific dipeptidyl aminopeptidases. Curr. Med. Chem. *12*, 971–98.

Barret, A.J., Rawlings, N.D. and Woessner (eds) (2004) Handbook of proteolytic enzymes. Elsevier Academic Press. Oxford, UK.

Bauvois, B. and Dauzonne, D. (2006) Aminopeptidase-N/CD13 (EC 3.4.11.2) inhibitors: chemistry, biological evaluations, and therapeutic prospects. Med. Res. Rev. *26*, 88–130.

Bazan, J.F., Weaver, L.H., Roderick, S.L., Huber, R. and Matthews, B.W. (1994) Sequence and structure comparison suggest that methionine aminopeptidase, prolidase, aminopeptidase P, and creatinase share a common fold. Proc. Natl. Acad. Sci. U S A. *91*, 2473–2477.

Beninga, J., Rock, K.L. and Goldberg, A.L. (1998) Interferon-gamma can stimulate post-proteasomal trimming of the N terminus of an antigenic peptide by inducing leucine aminopeptidase. J. Biol. Chem. *273*, 18734–18742.

Brown, N.J. (2005) Biological markers and diagnostic test for angiotensin converting enzyme inhibitor and vasopeptidase inhibitor-associated angioedema. Patent Application US6887679.

Burley, S.K., David, P.R., Taylor, A. and Lipscomb, W.N. (1990) Molecular structure of leucine aminopeptidase at 2.7-A resolution. Proc. Natl. Acad. Sci. U S A. *87*, 6878–6882.

Cappiello, M., Alterio, V., Amodeo, P., Del Corso, A., Scaloni. A., Pedone. C., Moschini, R., De Donatis, G.M., De Simone, G. and Mura, U. (2006) Metal ion substitution in the catalytic site greatly affects the binding of sulfhydryl-containing compounds to leucyl aminopeptidase. Biochem. *45*, 3226–3234.

Chen, S., Vetro, J.A. and Chang, T.H. (2002) The specificity *in vivo* of two methionine aminopeptidases in *Saccharomyces cerevisiae*. Arch. Biochem. Biophys. *398*, 87–93.

Chevalet, L., Souppe, J., De Leseleuc, J., Brunet, J. and Warmerdam, M.J. (2001) *Aspergillus niger* aminopeptidase compositions for making bread doughs and cheese. US6271013.

Chich, J.F., Rigolet, P., Nardi. M., Gripon, J.C., Ribadeau-Dumas. B. and Brunie. S. (1995) Purification, crystallization, and preliminary X-ray analysis of PepX, an X-prolyl dipeptidyl aminopeptidase from *Lactococcus lactis*. Proteins *23*, 278–281.

Christensen, J.E., Dudley, E.G., Pederson, J.A. and Steel, J.L. (1999) Peptidases and amino acid catabolism in lactic acid bacteria. Antonie van Leeuwenhoek *76*, 217–246.

Cogolludo, A., Perez-Vizcaino, F. and Tamargo, J. (2005) New insights in the pharmacological therapy of arterial hypertension. Curr. Opin. Nephrol. Hypertens. *14*, 423–427.

Cottrell, G.S., Hooper, N.M. and Turner, A.J. (2000) Cloning, expression, and characterization of human cytosolic aminopeptidase P: a single manganese(II)-dependent enzyme. Biochemistry *39*, 15121–15128.

Cunningham, D.F. and O'Connor, B. (1997) Proline specific peptidases. Biochim. Biophys. Acta *1343*, 160–186.

Di Cagno, R., De Angelis, M., Auricchio, S., Greco, L., Clarke, C., De Vincenzi, M., Giovannini, C., D'Archivio, M., Landolfo, F., Parrilli, G., Minervini, F., Arendt, E. and Gobbetti, M. (2004) Sourdough bread made from wheat and nontoxic flours and started with selected lactobacilli is tolerated in celiac sprue patients. Appl. Environ. Microbiol. *70*, 1088–1096.

Engel, M., Hoffmann, T., Wagner, L., Wermann, M., Heiser, U., Kiefersauer, R., Huber, R., Bode, W., Demuth, H.U. and Brandstetter, H. (2003) The crystal structure of dipeptidyl peptidase IV (CD26) reveals its functional regulation and enzymatic mechanism. Proc. Natl. Acad. Sci. U S A. *100*, 5063–5068.

Engel, C.K., Pirard, B., Schimanski, S., Kirsch, R., Habermann, J., Klingler, O., Schlotte, V., Weithmann, K.U. and Wendt K.U. (2005) Structural basis for the highly selective inhibition of MMP-13. Chem. Biol. *12*, 181–189.

FitzGerald, R.J. and O'Cuinn, G. (2006) Enzymatic debittering of food protein hydrolysates. Biotechnol. Adv. *24*, 234–237.

Fournie-Zaluski, M.C., Fassot, C., Valentin, B., Djordjijevic, D., Reaux-Le Goazigo, A., Corvol, P., Roques, B.P, and Llorens-Cortes, C. (2004) Brain renin-angiotensin system blockade by systemically active aminopeptidase A inhibitors: a potential treatment of salt-dependent hypertension. Proc. Natl. Acad. Sci. U S A. *101*, 7775–7780.

Gilboa, R., Spungin-Bialik, A., Wohlfahrt, G., Schomburg, D., Blumberg, S. and Shoham, G. (2001) Interactions of *Streptomyces griseus* aminopeptidase with amino acid reaction products and their implications toward a catalytic mechanism. Proteins *44*, 490–504.

Gonzales, T. and Robert-Baudouy, J. (1996) Bacterial aminopeptidases: properties and functions. FEMS Microbiol. *18*, 319–344.

Hauser, F., Strassner, J. and Schaller, A. (2001) Cloning, expression, and characterization of tomato (*Lycopersicon esculentum*) aminopeptidase P. J. Biol. Chem. *276*, 31732–31737.

Holz, R.C., Bzymek, K.P. and Swierczek, S.I. (2003) Co-catalytic metallopeptidases as pharmaceutical targets. Curr. Opin. Chem. Biol. *7*, 197–206.

l'Anson, K.J., Movahedi. S., Griffin. H.G., Gasson. M.J. and Mulholland, F. (1995) A non-essential glutamyl aminopeptidase is required for optimal growth of *Lactococcus lactis* MG1363 in milk. Microbiol. *141*, 2873–2881.

Inguimbert, N., Coric, P., Dhotel, H., Bonnard, E., Llorens-Cortes, C., Mota, N., Fournie-Zaluski, M.C. and Roques, B.P. (2005) Synthesis and *in vitro* activities of new non-peptidic APA inhibitors. J. Pept. Res. *65*, 175–188.

Jones, E.W. (1991) Three proteolytic systems in the yeast *Saccharomyces cerevisiae*. J. Biol. Chem. *266*, 7963–7966.

Joshua-Tor, L., Xu, H.E., Johnston, S.A. and Rees, D.C. (1995) Crystal structure of a conserved protease that binds DNA: the bleomycin hydrolase, Gal6. Science *269*, 945–950.

Kim, H., Burley, S.K. and Lipscomb, W.N. (1993) Re-refinement of the X-ray crystal structure of bovine lens leucine aminopeptidase complexed with bestatin. J. Mol. Biol. *230*, 722–724.

Kohno, H., Kanda, S. and Kanno, T. (1986) Immunoaffinity purification and characterization of leucine aminopeptidase from human liver. J. Biol. Chem. *261*, 10744–10748.

Kunji, E.R.S., Mierau, I., Hagting, A., Poolman, B. and Konings, W. N. (1996) The proteolytic systems of lactic acid bacteria. Antonie van Leeuwenhoek. *70*, 187–221.

Leclerc, P.L., Gauthier, S.F., Bachelard, H., Santure, M. and Roy, D. (2001) Antihypertensive activity of casein-enriched milk fermented by *Lactobacillus helveticus*. Int. Dairy J. *12*, 995–1004.

Leiting, B., Pryor, K.D., Wu, J.K., Marsilio, F., Patel, R.A., Craik, C.S., Ellman, J.A., Cummings, R.T. and Thornberry, N.A. (2003) Catalytic properties and inhibition of proline-specific dipeptidyl peptidases II, IV and VII. Biochem. J. *371*, 525–532.

Linderstrom-Lang, K. (1929) Über Darmereosin [Concerning intestinal erepsin]. Z. Physiol. Chem. *182*, 151–174.

Lipscomb, W.N. and Sträter, N. (1996) Recent Advances in Zinc Enzymology. Chem. Rev. *96*, 2375–2434.

Lowther, W.T. and Matthews, B.W. (2002) Metalloaminopeptidases: common functional themes in disparate structural surroundings. Chem. Rev. *102*, 4581–4608.

McDonald, J.K. and Barret, A.J. (eds.) (1986) Mammalian proteases: a glossary and bibliography. Vol 2. Exopeptidases. Academic Press, London.

Mata, L., Gripon, J.C. and Mistou, M.Y. (1999) Deletion of the four C-terminal residues of PepC converts an aminopeptidase into an oligopeptidase. Protein Eng. *12*, 681–686.

Matsushima, M., Inoue, H., Ichinose, M., Tsukada, S., Miki, K., Kurokawa, K., Takahashi, T. and Takahashi, K. (1991) The nucleotide and deduced amino acid sequences of porcine liver proline-beta-naphthylamidase. Evidence for the identity with carboxylesterase. FEBS Lett. *293*, 37–41.

Meisel, H. (2004) Multifunctional peptides encrypted in milk proteins. Biofactors *21*, 55–61.

Mest, H.J. (2006) Dipeptidyl peptidase-IV inhibitors can restore glucose homeostasis in type 2 diabetics via incretin enhancement. Curr. Opin. Investig. Drugs *7*, 338–343.

Meyer-Barton, E., Klein, J.R., Henrich, B. and Plapp, R. (1994) X-prolyl-dipeptidyl peptidase from *Lactobacillus delbrueckii* ssp. *lactis*, nucleic acids coding for the same and its use in fermented foodstuffs preparation process. Patent Application WO9416082.

Mistou, M.Y., Rigolet, P., Chapot-Chartier, M.P,, Nardi, M., Gripon, J.C, and Brunie, S. (1994) Crystallization and preliminary X-ray analysis of PepC, a thiol aminopeptidase from *Lactoccocus lactis* homologous to bleomycin hydrolase. J. Mol. Biol. *237*, 160–162.

Nampoothiri, K.M., Nagy, V., Kovacs, K., Szakacs, G. and Pandey, A. (2005) l-leucine aminopeptidase production by filamentous *Aspergillus* fungi. Lett. Appl. Microbiol. *41*, 498–504.

O'Farrell, P.A., Gonzalez, F., Zheng, W., Johnston, S.A. and Joshua-Tor, L. (1999) Crystal structure of human bleomycin hydrolase, a self-compartmentalizing cysteine protease. Structure *7*, 619–627.

Raksakulthai, R. and Haard, N.F. (2003) Exopeptidases and their application to reduce bitterness in food: a review. Crit. Rev. Food Sci. Nutr. *43*, 401–445.

Rigolet, P., Xi, X.G., Rety, S. and Chich, J.F. (2005) The structural comparison of the bacterial PepX and human DPP-IV reveals sites for the design of inhibitors of PepX activity. FEBS J. *272*, 2050–2059.

Roderick, S.L. and Matthews, B.W. (1993) Structure of the cobalt dependent methionine aminopeptidase from *Escherichia coli:* a new type of proteolytic enzyme. Biochem. *32*, 3907–3912.

Ross, S., Giglione, C., Pierre, M., Espagne, C. and Meinnel, T. (2005) Functional and developmental impact of cytosolic protein N-terminal methionine excision in Arabidopsis. Plant Physiol. *137*, 623–637.

Sanderink, G.J., Artur, Y. and Siest, G. (1988) Human aminopeptidases: a review of the literature. J. Clin. Chem. Biochem. *26*, 795–807.

Sanz, Y. and Toldrá, F. (1997) Purification and characterization of an aminopeptidase from *Lactobacillus sakei* . J. Agric. Food Chem. *45*, 1552–1558.

Sanz, Y. and Toldrá, F. (2002) Purification and characterization of an arginine aminopeptidase from *Lactobacillus sakei.* Appl. Environ. Microbiol. *68*, 1980–1987.

Sanz, Y., Sentandreu, M.A. and Toldrá, F. (2002) Role of muscle and bacterial peptidases in meat fermentation. In Research advances in the quality of meat and meat products. Toldrá F. (ed.). pp 143-155. Research Signpost, Trivandrum, ISBN: 81-7736-125-2, India.

Savijoki, K., Ingmer. H. and Varmanen, P. (2006) Proteolytic systems of lactic acid bacteria. Appl. Microbiol. Biotechnol. April 21; [Epub ahead of print]

Scharf, U., Stolz, P., Huscroft, S.C. and Schmidt-Hahn, K. (2006) Use of aminopeptidase in dough, doughs and bread improvers comprising aminopeptidase. Patent Application. WO2006009447.

Schiffmann, R., Neugebauer, A. and Klein, C.D. (2006) Metal-mediated inhibition of Escherichia coli methionine aminopeptidase: structure-activity relationships and development of a novel scoring function for metal-ligand interactions. J. Med. Chem. *49*, 511–522.

Scornik, O.A. and Botbol, V. (2001) Bestatin as an experimental tool in mammals. Curr. Drug Metab. *2*, 67–85.

Selvakumar, P., Lakshmikuttyamma, A., Dimmock, J.R. and Sharma, R.K. (2005) Methionine aminopeptidase 2 and cancer. Biochim. Biophys. Acta *1765*, 148–154.

Seppo, L., Jauhiainen, T., Pousa, T. and Korpela, R. (2003) A fermented milk high in bioactive peptides has a blood pressure-lowering effect in hypertensive subjects. Am. J. Clin. Nutr. *77*, 326–330.

Shigeri, Y., Tsujimoto, Y., Yutaka, M. and Kunihiko, W. (2005) Method for producing degraded material of collagen. Patent Application. JP2005245285.

Shimizu, T., Tani, K., Hase, K., Ogawa, H., Huang, L., Shinomiya, F. and Sone, S. (2002) CD13/aminopeptidase N-induced lymphocyte involvement in inflamed joints of patients with rheumatoid arthritis. Arthritis Rheum. *46*, 2330–2338.

Sträter N. and Lipscomb, W.N. (1995) Two-metal ion mechanism of bovine lens leucine aminopeptidase: active site solvent structure and binding mode of L-leucinal, a gem-diolate transition state analogue, by X-ray crystallography. Biochem. *34*, 14792–14800.

Sträter, N., Sherratt, D.J. and Colloms, S.D. (1999) X-ray structure of aminopeptidase A from *Escherichia coli* and a model for the nucleoprotein complex in Xer site-specific recombination. EMBO J. *18*, 4513–4522.

Suchiibun, S.Y. K., Maikuru, A.M. and Daramu, B.B. (1993) Flavor reinforcing method for natural beef concentrated liquid. Patent Application EP 0505733.

Tan, P.S.T., Poolman, B. and Konings, W.N. (1993) Proteolytic enzymes of *Lactococcus lactis*. J. Dairy Sci. *60*, 269–286.

Thunnissen, M.M.G.M., Nordlund, P. and Haeggström, J.Z. (2001) Crystal structure of human leukotriene A_4 hydrolase, a bifunctional enzyme in inflammation. Nat. Struct. Biol. *8*, 131–135.

Turner, A.J. (2004) Membrane alanyl aminopeptidase. In Handbook of proteolytic enzymes, Barret, A.J., Rawlings, N. D. and Woessner (eds), Vol 1 pp.289–294. Elsevier Academic Press. Oxford, UK.

Tu, C.J., Park, S-Y. and Walling, L.L. (2003) Isolation and characterization of the neutral leucine aminopeptidase (LapN) of tomato. Plant Physiol. *132*, 243–255.

Yamauchi, Y., Ejiri, Y. and Tanaka, K. (2001) Purification of an aminopeptidase preferentially releasing N-terminal alanine from cucumber leaves and its identification as a plant aminopeptidase N. Biosci. Biotechnol. .Biochem. *65*, 2802–2805.

Yaron, A. and Naider, F. (1993) Proline-dependent structural and biological properties of peptides and proteins. Crit. Rev. Biochem. Mol. Biol. *28*, 31–81.

Yeh, J.R., Mohan, R. and Crews, C.M. (2000) The antiangiogenic agent TNP-470 requires p53 and p21CIP/WAF for endothelial cell growth arrest. Proc. Natl. Acad. Sci. U S A. *97*, 12782–12787.

Yoshimoto, T., Orawski, A.T. and Simmons, W.H. (1994) Substrate specificity of aminopeptidase P from Escherichia coli: comparison with membrane-bound forms from rat and bovine lung. Arch. Biochem. Biophys. *311*, 28–34.

Yoshimoto, T., Kabashima, T., Uchikawa, K., Inoue, T., Tanaka, N., Nakamura, K.T., Tsuru, M. and Ito, K. (1999) Crystal structure of prolyl aminopeptidase from *Serratia marcescens*. J. Biochem. (Tokyo) *126*, 559–565.

Zheng, W., Johnston, S.A. and Joshua-Tor, L. (1998) The unusual active site of Gal6/bleomycin hydrolase can act as a carboxypeptidase, aminopeptidase, and peptide ligase. Cell *93*, 103–109.

Zhong, H. and Bowen, J.P. (2006) Antiangiogenesis drug design: multiple pathways targeting tumor vasculature. Curr. Med. Chem. *13*, 849–862.

SECTION C

LIPASES

CHAPTER 16

LIPASES: MOLECULAR STRUCTURE AND FUNCTION

MARINA LOTTI* AND LILIA ALBERGHINA

Dept. of Biotechnology and Biosciences, University of Milano-Bicocca, Milano, Italy
**marina.lotti@unimib.it*

1. INTRODUCTION

Early reports on the production by both bacterial and eukaryotic cells of enzymes able to degrade lipid substrates date to over a century ago. Since then, research on lipolytic enzymes – that includes lipases, esterases, phospholipases – has been driven by their central roles in lipid metabolism and in signal transduction. Lipases are generally versatile enzymes that accept a broad range of substrates (*i.e.* aliphatic, alicyclic, bicyclic and aromatic esters, thioesters, activated amines) whilst maintaining high regio-, chemo- and enantioselectivity. The stability of most lipases in organic solvents paves the way for their exploitation in organic synthesis: in esterification, transesterification, aminolysis and oximolysis reactions (Drauz and Waldman, 1995). Such properties make lipases key players in the industrial enzyme sector (Schmid and Verger, 1998; Bornscheuer, 2000; Kirk *et al*., 2002; Jaeger and Eggert, 2002; Gupta *et al*., 2004).

In this chapter we review the fundamental knowledge available on lipases, with particular emphasis on the relationship between the sequence, structure and function of those most commonly used in industrial processes. On the basis of this knowledge, novel and improved lipases may be generated, able to meet the requirements for robustness, selectivity and catalytic performances posed by modern biocatalysis.

2. BIOCHEMISTRY, FUNCTION AND EXPRESSION

2.1. Lipases Versus Esterases. The Concept of Interfacial Activation

Lipases are hydrolases and exert their activity on the carboxyl ester bonds of triacylglycerols and other substrates. Their natural substrates are insoluble lipid compounds prone to aggregation in aqueous solution. As early as 1958 Sarda

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 263–281

and Desnuelle described the sharp increase in lipase activity at substrate concentrations exceeding their solubility threshold as being the major difference to the esterases, enzymes active on ester bonds of soluble molecules that follow classical Michaelis- Menten kinetics (Sarda and Desnuelle, 1958). Since then the formation of an interface between aggregated substrates and the aqueous solution has been recognized as necessary for the activation of lipases (sometimes also referred to as "interfacial enzymes"). This behaviour – known as interfacial activation – found a structural rationale some years later when the first three-dimensional structures of lipase enzymes were elucidated (Winkler *et al*., 1990; Brady *et al*., 1990). These studies revealed that the enzyme active sites are shielded from the solvent by a mobile structure, the "lid" or "flap", that has to be displaced upon interaction with the substrate/water interface in order to yield an active enzyme conformation with the catalytic centre accessible to substrates. The crystal structures of lipases alone or in complexes with transition state analogues facilitated the elucidation of the conformational changes involved in the transition from the inactive closed lid conformation to the active one when the lid is open (Grochulski *et al*., 1994; Brzozowski *et al*., 1991). The mechanics of lid opening may vary between enzymes but in all cases leads to the creation of an open, accessible active site and a large hydrophobic lipid binding site (Fig. 1). In several lipases lid opening is also responsible for the formation of the so-called "oxyanion hole" which is involved in the stabilization of the reaction intermediates (see later). However, the classification of a lipolytic enzyme as being a *true* lipase (EC 3.1.1.3) on the basis of its activation at the interface and the presence of a lid structure does not hold in a number of cases. Lipases without a lid or with a lid but no interfacial activation have been described (Verger, 1997). To date, the broader definition of a lipase as a carboxylesterase catalysing hydrolysis and synthesis of long-chain acylglycerols is generally accepted and seems to be adequate to describe all known lipases. It specifically refers to

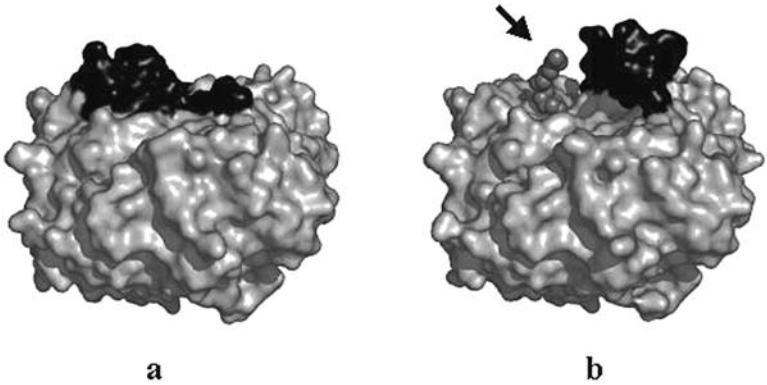


Figure 1. Lipase from *Candida rugosa* represented in the closed (a) and open conformation (b) with the lid depicted in black. In the active conformation (b) the enzyme active site is accessible to substrates here represented by an inhibitor (dark grey) and highlighted by the arrow

the behaviour of enzymes on insoluble substrates but it has to be recalled the most lipases are active also on soluble esterase substrates.

In low-water conditions, the reverse synthetic reaction is favoured, leading to esterification, alcoholysis and acidolysis. Such synthetic ability, along with the tolerance of several lipases to organic solvents (Zaks and Klibanow, 1984), is extensively exploited in organic synthesis (dealt with in depth in other chapters of this book).

2.2. General Molecular and Biochemical Features

Lipases are ubiquitous enzymes present in all types of living organisms. In eukaryotes they may be confined within an organelle (*i.e.* the lysosome), or they can be found in the spaces outside cells and play roles in the metabolism, absorption and transport of lipids. In lower eukaryotes and bacteria lipases can be either intracellular or be secreted in order to degrade lipid substrates present in the environment, and in some pathogenic organisms (*Candida albicans*, *Staphylococcus* and *Pseudomonas* species, *Helicobacter pylori*) they can even act as virulence factors. Enzymes from bacteria and fungi have the greatest potential as industrial biocatalysts since they are usually robust, easy to produce by fermentation and easy to recover from the culture broth. As a consequence, a large number of microbial enzymes can be obtained from commercial producers. Most bacterial lipases are sourced from *Pseudomonas, Burkholderia, Alcaligenes, Acinetobacter, Bacillus* and *Chromobacterium* species; widely used fungal lipases are produced by *Candida, Humicola, Penicillium, Yarrowia, Mucor, Rhizopus* and *Aspergillus* sp. Among the lipases from higher eukaryotes, porcine pancreatic lipase has been in use for several years as a technical enzyme. Other mammalian lipases are of medical interest as possible drug targets in the treatment of metabolic diseases or for direct development as drugs (Müller and Petry, 2004). In such cases, recombinant forms are favoured to overcome demanding purification protocols. For example, recombinant human gastric lipase is used in the treatment of pancreatic insufficiency caused by cystic fibrosis and pancreatitis. Plant enzymes, *e.g.* from papaya, pineapple, *Veronia*, *Euphorbia*, and in particular germinating seeds (castor bean, oil palm, oilseed rape), have interesting applications in biocatalysis as they display unusual fatty acid selectivities (Mukherijee and Hills, 1994). Such diversity in origin, cellular localization and function is reflected in an astonishing degree of biochemical variability since lipases from different organisms, or even isoenzymes produced by the same organism, may vary greatly in molecular mass, pH and temperature optima, post-translational modifications, and substrate and reaction specificities. This extensive variation is of importance to biotechnology as a potential source of biocatalysts endowed with a wide range of optima and specificities that can adapt to various process conditions. Attempts to broaden the biocatalytic power of the available lipases are taking a number of routes including the search for novel enzymes produced by organisms adapted to unusual habitats, the metagenomic approach, and rational and random mutagenesis of known enzymes.

Molecular masses of known lipases range from less than 20 kDa as in the case of the small lid-less lipolytic enzymes lipase A from *Bacillus subtilis* and cutinase from *Fusarium solani pisi*, to about 60kDa for the larger fungal lipases (*i.e. Geothricum candidum* lipase). In spite of this, almost all lipases share a common architecture and are structured in a single protein domain. Exceptions are found in lipases from higher eukaryotes where complex functions, *i.e.* interaction with other molecules and regulation, are attained through additional structural modules.

The range of temperature optima observed is wide, generally falling between 30 °C and 60 °C. However, this concerns lipases obtained from conventional sources. More recently the search for enzymes from extremophiles, *i.e.* organisms adapted to life in extreme environments, has enriched the spectrum with lipases with T_{opt} over 70 °C (*i.e. Bacillus thermocatenulatus* lipase) or those endowed with high activity at low temperature as is the case for enzymes produced by Antarctic bacteria, *i.e.* from *Pseudomonas* and *Moraxella* sp. Such extreme and unusual features open the possibility to apply these enzymes in their wild type form without the need for engineering approaches to adapt them for use in reactions carried out at high temperatures or, conversely low temperature processes such as that of detergents (low temperature washes) or in food processing (Demirhian *et al*, 2001). Most lipases used in biocatalysis have neutral or alkaline pH optima, in some cases up to or beyond pH 9.0 (*Pseudomonas* and *Bacillus* lipases). Less common are acidic lipases active at pH as low as ca. 3.0. Interestingly, some lipases from *Bacillus* sp. are active over a broad pH range (Gupta *et al*, 2004).

2.3. Control of Lipase Production

Lipases are involved in specific metabolic processes hence the expression of the genes encoding them is tightly regulated. The occurrence of these regulatory mechanisms has to be taken into particular account during the production of industrial lipases by fermentation, *i.e.* when dealing with bacterial or fungal producers. Expression of lipolytic proteins is often inducible and can be modulated by several parameters. Among them the carbon and nitrogen source provided during fermentation are of particular importance, as is the addition of compounds that can act as inducers, for example, fatty acids, Tweens, olive oil. Physiological parameters set during the fermentation protocol, such as the pH of the medium, temperature and oxygen supply also play roles (Gupta, 2004) since the production of lipases can be dependent on the growth phase of the culture as has been shown for *Streptomyces* and *Staphylococcus* strains as well as in *Pseudomonas aeruginosa* (Jaeger *et al*, 1999). Knowledge about the regulation of gene expression is of particular relevance in several known cases where the source organism produces lipase isoenzymes, *i.e.* related proteins encoded by a family of paralogous genes. Usually protein isoforms are closely related in sequence and biochemical features, but not identical, and differences can be relevant from a catalytic point of view. Good examples are provided by fungal strains, as for example the asporogenic yeast *Candida rugosa* which produces at least 7 proteins differing in substrate specificity,

glycosylation, temperature and pH stability (Lotti *et al.*, 1993; Lopez *et al.*, 2004), *Yarrowia lipolytica* (Fickers *et al.*, 2005) and the opportunistic pathogen *Candida albicans* which has at least ten lipase proteins (Hube *et al.*, 2000). In such organisms the expression of isoenzymes can be subjected to complex control mechanisms. This issue has been studied in detail for the *Candida rugosa* lipases, some of which are constitutively expressed whilst others are induced by substrates present in the medium (Lotti *et al.*, 1998; Lee *et al.*, 1999). Whereas the availability of related and complementary enzymatic activities has obvious metabolic advantages for the producing strains, it can lead to enzymatic preparation of poorly reproducible composition and/or catalytic performance (Lopez *et al.*, 2004).

2.4. Occurrence and Functional Relevance of Post-Translational Modifications

Eukaryotic lipases are often glycosylated. The role of sugar chains in the activity, stability and secretion of a number of lipases has been investigated in depth using mutant proteins lacking glycosylation sites. However, determining the functional role of oligosaccharides is not always straightforward. In most cases they affect protein solubility and, as a consequence, the folding and/or the secretion of the enzyme (Miller *et al.*, 2004). Nevertheless, some specific functional roles have been elucidated. A clear role for asparagine-linked sugars has been pointed out in enzymes belonging to the acid lipase family, characterized by stability and activity under low pH conditions. Human gastric lipase (HGL) for example, which initiates the digestion of triglycerides in the stomach, is a highly glycosylated protein (up to 15% of the protein mass) with four potential N-glycosylation sites. The activity of deglycosylated recombinant HGL is affected to different extents depending on the number of sugar chains removed, but the most evident impact of deglycosylation is the increased susceptibility to pepsin degradation in acidic conditions shown by the deglycosylated enzyme (Wicker-Planquart *et al.*, 1999). An active role in enzyme activation, *i.e.* in lid opening, has been shown in two fungal lipases. Removal of an asparagine residue strictly conserved in the *Candida rugosa* lipase family resulted in a dramatic drop in enzyme activity whereas deglycosylation at other locations impacted to a much lower extent on activity (Brocca *et al.*, 2000). In this case, crystallographic analysis of the enzyme in the open and closed forms suggested that this sugar chain contributes to the stabilization of the open active form by interacting with the inner surface of the open lid (Grochulski *et al.*, 1994; Fig. 2a). The second example concerns a non-glycosylated mutant of *Thermomyces lanuginosa* lipase which displays lower binding affinity to phospholipid liposomes. This behaviour is suggested to affect the dynamics of lid movement and, as a consequence, the binding of the enzyme to the interface (Peters *et al.*, 2002). These and a number of other reports clearly indicate that the glycosylation ability of the host has to be carefully considered in heterologous expression of lipases, as sugar chains appear to impact on several issues of lipase functionality.

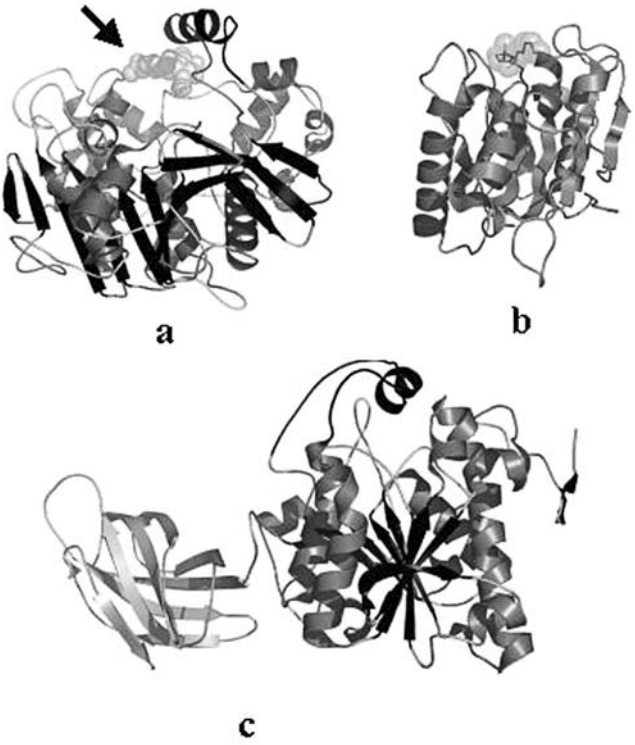


Figure 2. Variation on the α/β hydrolase fold design in lipases of different complexity: (a) the *Candida rugosa* enzyme structure where the arrow marks the oligosaccharide chain linked to Asn 351, (b) the mini-lipase from *Bacillus subtilis* distinguished by the lack of a lid structure and (c) the human pancreatic lipase with the colipase binding domain on the left side

Rare and so far unique to lipases subjected to hormonal regulation, is reversible phosphorylation. Hormone-sensitive lipase (HSL) is responsible for the mobilization of fatty acids in adipose tissue in response to hormonal stimuli and is regulated by phosphorylation by a number of protein kinases, in particular by cAMP-dependent protein kinase A. Four serine residues have been identified as kinase targets. The mechanism leading to HSL phosphorylation-mediated activation seems to involve not just conformational changes but also translocation of the protein from the cytosol to lipid droplets (for a recent review see Yeaman, 2004).

2.5. Specificity (Selectivity) of Lipase-Catalysed Reactions

The potential of lipases as biocatalysts relies on their sophisticated selectivity and specificity which permits the fine tuning of reactions. Specificity or selectivity can concern regioselectivity, *i.e.* the position in the substrate molecules of the ester bonds hydrolysed or formed; chemo-selectivity, *i.e.* the nature of the substrate

recognized; and stereoselectivity. One field of biocatalysis where such properties are successfully exploited is the modification of triglycerides where three features are relevant: i) regioselectivity *i.e.* the position of the fatty acid on the glycerol backbone; ii) fatty acid specificity concerning *i.e.* the length or unsaturation of the chain; iii) the class of acylglycerols, *i.e.* mono-, di- or triglycerides. Most known lipases are 1, 3 regiospecific with activity on the primary alcohol positions whereas only a few are able to recognize also the sn-2 position allowing for the complete hydrolysis of triglycerides to free fatty acids. Concerning fatty acid selectivity, lipases are able to convert esters of medium to long chain (C4 to C18, rarely up to C22) but with different efficiencies. Even isoforms of the same enzyme can differ in this property. This is the case for the isoforms of *Candida rugosa* lipase where isoform 1 acts mainly on medium chain (C8-C10) substrates, isoform 3 on short-chain soluble substrates, and isoforms 2 and 4 on long-chain molecules (C16-C18). Some lipases display unusual preferences towards unsaturated fatty acids. Worthy of mention in this regard are one isoform of *Geotrichum candidum* lipase selective for *cis* (Δ-9) unsaturated substrates, pancreatic lipase and some microbial lipases active on long-chain polyunsaturated substrates (PUFA), and others (from guinea pig, *S. hyicus*, *Rhizopus*) with phospholipase A1 activity. Lipolytic enzymes possessing different selectivities can therefore be used alone or in combination to obtain valuable products, such as structured triglycerides with improved nutritional value, cocoa butter substitutes, and oils enriched in PUFAs, as well as an impressive range of mono- di- and triacylglycerols, fatty acids, esters and intermediates (Bornscheuer, 2000). Another field where lipases find increasing application is in the regioselective acylation of polyfunctional molecules such as carbohydrates, amino acids and peptides - in particular in the protection/deprotection steps necessary for the generation of combinatorial libraries on carbohydrate scaffolds for the development of new drugs (Le *et al*, 2003). Another property of lipases of paramount importance for application in fine chemistry and drug and agrochemical production, is their stereoselectivity toward a broad range of substrates which facilitates reactions on prochiral substrates and the kinetic resolution of racemates. The use of lipases in such processes extends to prochiral and chiral alcohols, carboxylic acid esters, α and β-hydroxy acids, diesters, lactones, amines, diamines, amino-alcohols, α- and β-amino acid derivatives (Schmidt *et al*, 2001). Examples of industrial scale lipase-catalysed processes include the kinetic resolution of various amines and the production of an intermediate in the synthesis of DiltiazemTM (a calcium antagonist used to control high blood pressure) by *Serratia marcescens* lipase (Shibatani *et al*, 1990).

3. DIVERSITY AND CONSERVATION WITHIN LIPASES: SEQUENCES AND STRUCTURES

3.1. Primary Sequences and Sequence-Based Classification of Lipases

By the end of 2005 about 2000 non-redundant sequences of lipases and related enzymes were present in protein sequence databases. No specific sequence similarity

is shared by all known lipases. On the contrary, they appear to be astonishingly variable. In the Lipase Engineering Database (LED), lipases are grouped in 16 superfamilies and 39 homologous families (Fisher and Pleiss, 2003). The lone consensus shared by all of them is the pentapetide Gly-X-Ser-X-Gly (with rare cases where glycines are substituted by other small residues). This motif, which encloses the active site serine, is denominated in the PROSITE database (Hulo *et al.*, 2004) as PS00120 ([LIV]-{KG}-[LIVFY]-[LIVMST]-G-[HYWV]-S-{YAG}-G-[GSTAC]) and identifies a proteins as a lipase.

3.2. All Lipases Share a Common Structural Fold

Despite their variability in primary sequence, all lipases display the same structural architecture, the so-called α/β hydrolase fold, and have identical catalytic machineries. Such structural conservation is a very valuable tool helping in the classification of newly identified proteins even in the absence of clear sequence similarity. Moreover, it facilitates modelling approaches prior to protein engineering experiments. The original description of this fold was based on the comparison of the three-dimensional structures of hydrolases mostly unrelated in primary sequence and active on substrates very different in structure, one of which was a fungal lipase (Ollis *et al.*, 1992). All lipases whose 3D structures were later solved were found to be members of this fold family. The design of the canonical α/β hydrolase fold is based on a central, mostly parallel β-sheet of eight strands with the only strand (β2) antiparallel. Strands β3 to β8 are connected by α-helices packed on both sides of the β-sheet. Variations from the canonical fold can affect the number of β-strands, the presence of insertions, and the architecture of the substrate binding subdomains (Fig. 2). Lipases of known 3D structure are currently classified by the SCOP database (Murzin *et al.*, 1995) into 7 families based on the elements of the basic fold that they contain: acetylcholinesterase-like, gastric lipase, lipase, fungal lipase, bacterial lipase, pancreatic lipase N-terminal domain, and cutinase-like. The small bacterial lipase A from *Bacillus subtilis* has been defined as a "minimal α/β hydrolase fold protein" as it only contains a six-stranded parallel β-sheet flanked by five α-helices (van Pouderoyen *et al.*, 2001). Additional domains can be added to this basic architecture, *i.e.* in enzymes involved in protein-protein or protein-lipid interactions or those subjected to regulation such as pancreatic lipase and hormone-sensitive lipase.

In α/β hydrolases the active site consists of a catalytic triad comprising a nucleophile, an acidic residue and a histidine, reminiscent of that of serine proteases but with a different order in the sequence: nucleophile-acid-histidine (Ollis *et al.*, 1992). The lipase catalytic triad is composed of serine, aspartate or glutamate and histidine, with the serine enclosed in the consensus motif previously mentioned which forms a sharp turn (the nucleophile elbow) in a strand-turn-helix motif in strand β5 which forces the nucleophile to adopt unusual main chain Φ and Ψ torsion angles. Due to its functional relevance, the nucleophile elbow is the most conserved feature

of the fold. Hydrolysis of the substrate follows a two-step mechanism. The nucleophilicity of the active serine is enhanced by transferring a proton to the catalytic histidine with the formation of an oxyanion that attacks the carbonyl carbon of the susceptible ester bond. A tetrahedral intermediate is formed carrying a negative charge on the carbonyl oxygen atom of the scissile bond and it is stabilized through hydrogen bonding to main-chain NH groups. Such residues build up the so-called oxyanion hole that in some lipases is preformed in the correct orientation, whereas in others it is positioned upon the opening of the lid structure. The proton on the histidine is then transferred to the ester oxygen of the bond that is cleaved and a covalent intermediate forms with the fatty acid from the substrate esterified to serine. The second step of the reaction is deacylation of the enzyme through a water molecule that hydrolyses the covalent intermediate. In this case, transfer of a proton from water to the active site serine produces a hydroxide ion that attacks the carbonyl carbon atom in the substrate–enzyme covalent intermediate. In addition the negatively charged tetrahedral intermediate is stabilized by hydrogen bonds to the oxyanion hole. Finally, histidine donates a proton to the oxygen atom of the active serine and the acyl component is released.

3.3. Complexity in Lipases From Eukaryotes: Modularity and Regulation

Some lipolytic enzymes active in eukaryotic cells are faced with demanding functions that require additional abilities, such as the interactions with lipids under unfavourable conditions, membranes, other molecules, and more rarely, regulation. The two best characterized examples are pancreatic lipase (PL) and hormone-sensitive lipase (HSL). Both enzymes are organized in modules with a catalytic domain with the functional and structural characteristics previously described, plus additional domains that confer other properties.

PL is composed of two domains connected by a flexible hinge, a large N-terminal catalytic domain and a β-sandwich C-terminal domain which is related to the peculiar physiological environment in which the enzyme has to be active. In the intestinal lumen dietary triglycerides are mixed with phospholipids, fatty acids, proteins and bile salts that act as emulsifiers. Bile salts would prevent PL from adsorbing to the lipid substrate were it not for the association with a small protein – colipase – that is co-secreted by the pancreas. Colipase is an amphiphilic protein able to anchor the lipase to the lipid interface and stabilize it in the active open conformation. Upon binding to the lipase C-terminal domain colipase exposes hydrophobic finger structures on the opposite site and brings the enzyme in contact with the interface. Colipase binding does not induce conformational changes in the lipase molecule but indirectly allows opening of the lid through contact with the interface. However the cofactor makes contact with the open lid and with it forms a large hydrophobic surface able to interact strongly with the lipid-water interface (van Tilbeurgh *et al*., 1992).

Hormone-sensitive lipase (HSL) is an intriguing enzyme whose complex functions are still not completely unravelled. Its major and best characterized activity is the hydrolysis of triacylglycerols stored in adipose tissue, the first and rate-limiting step in the mobilization of fatty acids. HSL is composed of two structural domains with the active site in the C-terminal module. Phosphorylation sites are located in an extra module that interrupts the sequence of the catalytic domain. In the tertiary structure this module protrudes from the core of the domain which can therefore assume the canonical α/β hydrolase fold. In addition, HSL contains an N-terminal domain involved in protein-protein and protein-lipid interactions, as the enzyme has to make contact with lipid droplets accumulated in tissues. The main interactor of this docking domain has been shown to be the fatty acid-binding protein (FABP) that facilitates the release of fatty acids and their intracellular diffusion (Jenkins-Kruchten *et al*., 2003). HSL, which is subjected to several levels of regulation including reversible phosphorylation, translocation and association with regulatory proteins, provides an interesting example showing that new properties can be introduced in a lipase without interfering with its fold and conformation (Yeaman, 2004).

4. DETERMINANTS OF LIPASE SPECIFICITY

Lipase selectivity has been studied from several points of view with the aim of understanding its molecular and conformational basis on the one hand, and to be able to modulate enzyme performances on the other. The molecular features of the enzyme, the chemical structure of the substrate and the reaction conditions are the three major factors affecting specificity. With regard to the latter, several studies have been devoted to assess the influence of the solvent, the quality of the substrate interface and the matrix used to immobilise the biocatalyst (Cernia and Palocci, 1997; Villeneuve *et al*., 2000). Medium engineering has explored the effects of different organic and non-conventional solvents and water activity conditions and, more recently, the influence of ionic liquids has been examined (Park and Kazlauskas, 2003). However, understanding the molecular basis of lipase selectivity is a prerequisite for modifying the properties of the enzyme, hence this has been investigated in depth during recent years making use of synergic and complementary approaches: i) X-ray analysis of lipases in complex with substrates or their analogues; ii) the generation of site-specific and random mutants; iii) modelling of the available experimental results to extrapolate general rules and acquire predictive capabilities. From such investigations two structural elements came into focus as being major determinants of lipase specificity: the substrate binding site and the lid.

4.1. The Substrate Binding Site

The active site in lipases is buried within the protein structure and substrate access to it is through a binding site located in a pocket on the top of the central β-sheet. Although lipases share the same structural fold their substrate binding regions are

considerably different in size, structure and physico-chemical features, in particular regarding the hydrophobicity of residues lining the pocket. The length, shape and hydrophobicity of the binding pocket has been related to chain length preference, obtaining good agreement with experimental results (Pleiss *et al*, 1998). Based on this information and X-ray determinations of the structures of complexes to substrate analogues, several mutant enzymes have been created by introducing bulkier or more hydrophilic residues at the entrance, along the walls and at the bottom of the binding pocket in lipases from *Mucor miehei*, *Rhizopus*, *Humicola lanuginosa* and *Candida rugosa* (see for example Klein *et al*, 1997; Schmitt *et al*, 2002). In most cases this results in a change in the relative activity toward ester or lipid substrates of different chain lengths. These results confirmed the central role of the substrate binding site and showed that specific shifts in selectivity can be planned based on the analysis of structural and docking data.

It has been more difficult to define general rules explaining the stereopreference of lipases toward chiral and prochiral substrates. This appears to depend on both the substrate structure and on the lipase used, and is strongly influenced by the reaction conditions (Ransac *et al*, 1990). Attempts to rationalize the structural bases of stereopreference aimed at the identification of the binding regions of the acyl and alcohol portions of substrates. This was approached by crystallographic analysis of complexes of lipases with transition state analogues of fast- and slow-reacting enantiomers. A detailed structural analysis of the binding of a lipid analogue to *Burkholderia cepacia* lipase led to the identification of four binding pockets for the substrate: the oxyanion hole and three pockets lined by hydrophobic amino acids that accommodate the *sn*-1, *sn*-2 and *sn*-3 fatty acid chains. A central role is played by hydrogen bonding between the ester oxygen atom of the *sn*-2 chain and the histidine of the active site, and the sn-2 pocket is identified as the major determinant of the enzyme's stereopreference (Lang *et al*, 1998). This is in good agreement with experiments pointing to the importance of the substituent at the *sn*-2 position of the substrate (Kovac *et al*, 2000), and found further support from site-directed mutagenesis performed on the residues lining the binding pockets. A general conclusion can be drawn from the studies reported in this section, *i.e.* that the size, shape and hydrophobicity/hydrophilicity of the various substrate binding pockets are key players in determining lipase enantio- and regio-preferences and are therefore obvious targets for mutagenesis aiming to improve/modify these properties. Based on rational design, the enantioselectivity of the *Candida antarctica* B lipase catalysed resolution of 1-chloro-2-octanol was improved from E=14 to 28 by a single amino acid exchange as predicted by molecular modelling (Roticci *et al*, 2001).

4.2. The Lid

Lipases occur in alternative conformational states stabilised by the interaction with water/substrate interfaces. In the closed conformation the lid covers the enzyme active site, making it inaccessible to the substrate molecules, whereas transition to

the open conformation opens the entrance of the catalytic tunnel. In recent years it has become clear that the function of this lid is not simply to act as a gate that regulates access to the active site. Lids are amphipathic structures: in the closed enzyme structure their hydrophilic side faces the solvent and the hydrophobic face is directed towards the protein core. As the enzyme shifts to the open conformation, the hydrophobic face becomes exposed and contributes to the formation of a larger hydrophobic surface and the substrate binding region (Fig. 1). Studies by several groups have pointed to the lid as being a major molecular determinant of lipase activity and selectivity. Thus, for example, two members of the lipase gene family, human pancreatic lipase and guinea pig pancreatic lipase-related protein 2 differ in specificity in that the former enzyme shows high activity only on triglycerides whereas the latter has additional phospholipase and galactolipase activities. The main structural difference between the two enzymes concerns the presence in the guinea pig protein of a lid of extremely reduced size (5 amino acids). Site-directed mutagenesis and the creation of chimeras with exchanged lids revealed the role of the lid domain in the selectivity towards triglycerides, phospholipids and galactolipids (Carrière *et al*., 1998). Other examples pointing to a crucial role of the lid in substrate selectivity are *Candida rugosa*, *Pseudomonas* and *Bacillus* lipases, among others. *Candida rugosa* produces isoenzymes of differing substrate specificities, of which only isoforms 2 and 3 hydrolyse cholesterol esters. Replacement of the lid of isoform 1, which is completely inactive on such substrates, was sufficient to improve activity on cholesteryl linoleate by 200 fold (Brocca *et al*., 2003). The lipase from *Pseudomonas fragi* is highly specific for short-chain substrates whereas closely related enzymes from *Pseudomonas* and *Burkholderia* sp. prefer medium- or long-chain substrates. Mutagenesis of specific residues of the lid produced a shift in chain length preference towards medium-chain molecules (Santarossa *et al*., 2005). Whether the effect on specificity can be directly attributed to the sequence of the lid structure is not always clear, and possible effects on the flexibility and conformation of this structure that might be of importance for enzyme-substrate interactions cannot be excluded. This region of the protein is therefore a good target for protein engineering, since the lid is a surface loop and is likely to tolerate amino acid substitutions, insertions and deletions easier than structures buried in the core of the protein (Eggert *et al*., 2004).

5. PERSPECTIVES FOR LIPASE RESEARCH

In recent years the importance of lipases as industrial catalysts has grown steadily, raising interest in finding new enzymes endowed with novel and often non-natural properties. It is well recognized that the catalytic ability and specificity of lipases can be considerably influenced by the experimental conditions and therefore methods to modulate catalytic behaviour through, for example, reaction engineering are exploited in several laboratories. However, direct manipulation of the biocatalyst appears to be the most straightforward approach. Several ways are open to

researchers, among which two are considered below: i) the search for novel enzymes from organisms adapted to unusual and poorly explored environments or organisms that cannot be cultured in the laboratory, and ii) engineering of already known enzymes by rational engineering or random mutagenesis.

5.1. Search for Novel Enzymes by Exploiting Biodiversity

Organisms exploited as enzyme producers represent just a tiny fraction of those existing in nature. Environments extreme for temperature, pH or salt concentration are sources of adapted organisms that often produce proteins with unusual properties (Demirhian, 2001). A large number of micro-organisms are not amenable to laboratory cultivation and therefore completely unexplored regarding their catalytic repertoire. The so-called metagenomic approach aims to isolate genes of interest from non-characterized samples without any need to cultivate and/or isolate them. It relies on the construction of gene libraries from samples directly taken from the environment (soil, water) enriched in the organisms/activities of interest followed by the screening of the library obtained (Henne *et al*, 2000; Voget *et al*, 2003). Additionally, the list of genomes completely sequenced is constantly growing and sequences available in public databases can be screened by bioinformatic methods to identify putative lipase genes that are then amplified by PCR (Kim *et al*, 2004).

5.2. Construction of New Enzymes by Protein Engineering

Several recombinant lipases have been expressed in bacterial, fungal, plant and insect systems and can therefore be subjected to mutagenesis. Rational protein design is applied to lipases whose 3D structure has been solved (Table 1) or can be modelled by homology. A large number of site-specific mutants or chimeric proteins has been generated with the purpose of addressing lipase stability and specificity (for a review see Svendsen, 2000). Several successful cases can be cited, among them the enhancement of the enantioselectivity of *Candida antarctica* B lipase in the resolution of 1-chloro-2-octanol by virtue of a single amino acid substitution, or the expansion of the range of secondary alcohols it accepts by rational redesign of the stereospecificity pocket (Roticci *et al*, 2001; Magnusson *et al*, 2005). An ambitious goal of protein engineering is to obtain so-called “enzyme promiscuity”, which refers to the ability of an enzyme to catalyse more than one chemical transformation, such as the formation of carbon-carbon bonds by *C. antarctica* lipase B (Kazlauskas, 2005).

Directed evolution relies on the generation of libraries of random mutants followed by selection of those variants with improved qualities on which further rounds of mutagenesis can be performed (Arnold and Georgiou, 2003). In recent years directed evolution has been applied to a large number of proteins and enzymes for the purpose of improving activity, stability and specificity. This is especially useful when structural data are not available for rational protein engineering or in those cases where the determinants of the required feature are complex, *e.g.* for

Table 1. Lipases of known 3D structure

Organism	Reference
Bacteria	
Burkholderia glumae	Noble *et al.,* 1993
Burkoholderia cepacia	Scharg *et al.,* 1997
Pseudomonas aeruginosa	Nardini *et al.,* 2000
Bacillus subtilis	van Pouderoyen *et al.,* 2001
Streptomyces exfoliatus	Wei *et al.,* 1998
Bacillus stearothermophilus	Tyndall *et al.,* 2002
Fungi	
Candida rugosa 1	Grochulski *et al.,* 1993
Candida rugosa 2	Mancheno *et al.,* 2003
Candida rugosa 3	Ghosh *et al.,* 1995
Thermomyces lanuginosa	Brzozowski *et al.,* 2000
Candida antarctica	Uppenberg *et al.,* 1994
Rhizopus niveus	Kohno *et al.,* 1996
Rhizomucor miehei	Brady *et al.,* 1990
Geothricum candidum	Scharg *et al.,* 1993
Penicillium camembertii	Derewenda *et al.,* 1994
Fusarium solani cutinase	Martinez *et al.,* 1992
Higher Eukaryotes	
Human pancreatic	van Tilbeurgh *et al.,* 1992
Horse pancreatic	Lombardo, 1989
Bile-salt activated	Terzyan *et al.,* 2000
Human gastric	Roussel *et al.,* 1999
Dog gastric	Roussel *et al.,* 2002
Rat pancreatic lipase-related pr2	Roussel *et al.,* 1998

Table 2. Recent examples of lipases modified by directed evolution

Organism	Modification	Reference
Pseudomonas aeruginosa	enantioselectivity	Liebton *et al.,* 2000
Pseudomonas aeruginosa	inversion of enantioselectivity	Zha *et al.,* 2001
Pseudomonas aeruginosa	range of substrate accepted	Reetz *et al.,* 2005
Pseudomonas aeruginosa	amidase activity	Fujii *et al.,* 2005
Bacillus thermocatenulatus	phospholipase activity	Kauffmann and Schmidt-Dannert, 2001
Bacillus subtilis	inversion of enantioselectivity	Funke *et al.,* 2003
Bacillus subtilis	thermostability	Acharya *et al.,* 2004
Bacillus subtilis	enantioselectivity	Eggert *et al.,* 2005
Burkholderia cepacia	inversion of enantioselectivity	Koga *et al.,* 2003
Candida antarctica B	activity and thermostability	Suen *et al.,* 2004
Candida antarctica B	enantioselectivity, secondary alcohols	Qian and Lutz, 2005
Acinetobacter sp	hydrolytic activity	Han *et al.,* 2004
Rhizopus oryzae	reaction specificity	Shibamoto *et al.,* 2004
Metagenome esterase	lipase activity	Reyes-Duarte *et al.,* 2005

enantioselectivity or stability (Table 2). The most extensive and successful example reported so far is the evolution of the enantioselectivity of a *Pseudomonas aeruginosa* lipase towards 2-methyldecanoate. Rounds of directed evolution followed by saturation mutagenesis on the positions identified as "hot spots" for selectivity enhanced selectivity from E=1 to E=50 and produced mutants with reverse stereopreference (Liebton *et al*, 2000; Zha *et al*, 2001). The same approach has been also applied for improving the phospholipase A1 activity of two bacterial lipases (Kauffmann and Schmidt-Dannert, 2001; van Kampen and Egmond, 2000).

6. CONCLUSIONS

Despite their broad diffusion in biotransformation reactions the use of lipases (and of most enzymes) in industrial processes is still limited by intrinsic weaknesses of the biological catalyst, in particular low stability under operational conditions and low activity or specificity on particular or non-natural substrates. In this chapter, the potential of lipases has been emphasized and attention has been drawn to recent developments that are expected to expand the natural abilities of these proteins. Knowledge of the molecular determinants of enzyme properties has accumulated allowing the rational choice or creation of the "right catalyst" for a given process. On the other hand, the cloning of genes encoding as yet unknown enzymes from non-conventional sources and the modification of those already available by a combination of molecular techniques are very promising as potential sources of novel catalysts with improved or completely new properties.

REFERENCES

Acharya, P., Rajakumara, E., Sankaranarayanan, R. and Rao, N.M. (2004) Structural basis of selection and thermostability of laboratory evolved *Bacillus subtilis* lipase. J. Mol. Biol. *341*(5), 1271–1281.

Arnold, F.H. and Georgiou, G. (2003) Directed Evolution vols 230 and 231. Humana Press, Totowa, New Jersey.

Bornscheuer, U.T. (ed.) Enzymes in lipid modification. Wiley VCH, 2000.

Bornscheuer, U.T., Bessler, C., Srinivas, R. and Krishna, S.H. (2002) Optimizing lipases and related enzymes for efficient application. Trends Biotechnol. *20*(10), 433–437.

Brady, L., Brzozowski, A.M., Derewenda, Z.S., Dodson, E., Dodson, G., Tolley, S., Turkenburg, J.P., Christiansen, L., Huge-Jensen, B., Norskov, L., Thim, L. and Menge, U. (1990) A serine protease triad forms the catalytic center of a triacylglycerol lipase. Nature *343*, 767–770.

Brocca, S., Persson, M., Wehtje, E., Adlercreutz, P., Alberghina, L. and Lotti, M. (2000) Mutants provide evidence of the importance of glycosydic chains in the activation of lipase 1 from *Candida rugosa*. Prot. Sci. *9*, 985–990.

Brocca, S., Secundo, F., Ossola, M., Alberghina, L., Carrea, G. and Lotti, M. (2003) Sequence of the lid affects activity and specificity of *Candida rugosa* lipase isoenzymes. Prot. Sci. *12*, 2312–2319.

Brzozowski, A.M., Derewenda, U., Derewenda, Z.S., Dodson, G.G., Lawson, D.M., Turkenburg, J.P., Bjorkling, F., Huge-Jensen, B., Patkar, S.A. and Thim, L. (1991) A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex. Nature *351*, 491–494.

Brzozowski, A.M., Savage, H., Verma, C.S., Turkenburg, J.P., Lawson, D.M., Svendsen, A. and Patkar, S. (2000) Structural origins of the interfacial activation in *Thermomyces (Humicola) lanuginosa* lipase. J. Biochem. *39*, 15071–15082.

Carrière, F., Withers-Martinez, C., van Tilbeurgh, H., Roussel, A., Cambillau, C. and Verger, R. (1998) Structural basis for the substrate selectivity of pancreatic lipases and some related proteins. Biochim. Biophys. Acta *1376*, 417–432.

Cernia, E.and Palocci, C. (1997) Lipases in supercritical fluids. Methods Enzymol. *286*, 495–508.

Demirhian, D.C., Moris-Varas, F. and Cassidy, C.S. (2001) Enzymes from extremophiles. Curr. Opinion Chem. Biol. *5*, 144–151.

Derewenda, U., Swenson, L., Green, R., Wei, Y., Dodson, G.G., Yamaguchi, S., Haas, M.J. and Derewenda, Z.S. (1994) An unusual buried polar cluster in a family of fungal lipases. Nat. Struct. Biol. *1*, 36–47.

Drauz, K. and Waldmann, H. (1995) Enzyme Catalysis in Organic Synthesis. Wiley-VCH Verlag, GmbH, Weinheim.

Eggert, T., Leggewie, C., Puls, M., Streit, W., van Pouderoyen, G., Dijkstra, B.W. and Jaeger, K.-E. (2004) Novel Biocatalysts by identification and design. Biocat. Biotrans. *22*, 139–144.

Eggert, T., Funke, S.A., Rao, N.M., Acharya, P., Krumm, H., Reetz, M.T. and Jaeger, K.-E. (2005) Multiplex-PCR-based recombination as a novel high-fidelity method for directed evolution ChemBiochem *6*(6), 1062–1067.

Fickers, P., Fudalej, F., Le Dall, M.T., Casaregola, S., Gaillardin, C., Thonart, P. and Nicaud, J.M. (2005) Identification and characterization of LIP7 and LIP8 genes encoding two extracellular triacylglycerol lipases in the yeast *Yarrowia lipolytica*. Fungal genetics and biology *42*(3), 264–274.

Fisher, M. and Pleiss, J. (2003) The lipase engineering database: a navigation and analysis tool for protein families. Nucl. Acid Res. *31*(1), 319–321.

Fujii, R., Nakagawa, Y., Hiratake, J., Sogabe, A. and Sakata, K. (2005) Directed evolution of *Pseudomonas aeruginosa* lipase for improved amide-hydrolyzing activity. Prot. Eng. Des. Sel. *18*(2), 93–101.

Funke, S.A., Eipper, A., Reetz, M.T., Otte, N., Thiel, W., Van Pouderoyen, G., Dijkstra, B.W., Jaeger, K.-E. and Eggert, T. (2003) Directed evolution of an enantioselective *Bacillus subtilis* lipase. Biocat. Biotransf. *21*(2), 67–73.

Ghosh, D., Wawrzak, Z., Pletnev, V.Z., Li, N.Y., Kaiser, R., Pangborn, W., Jornvall, H., Erman, M. and Duax, W.L. (1995) Structure of uncomplexed and linoleate-bound *Candida cylindracea* cholesterol esterase. Structure *3*, 279–288.

Grochulski, P., Li, Y., Schrag, J.D., Bouthillier, F., Smith, P., Harrison, D., Rubin, B. and Cygler, M. (1993) Insights into interfacial activation from an open structure of *Candida rugosa* lipase. J. Biol. Chem. *268*, 12843–12847.

Grochulski, P., Li, Y., Schrag, J.D. and Cygler, M. (1994) Two conformational states of *Candida rugosa* lipase. Prot. Sci. *3*, 82–91.

Gupta, R., Gupta, N. and Rathi, P. (2004) Bacterial lipases: An overview of production, purification and biochemical properties. Appl. Microbiol. Biotechnol. *64*, 763–781.

Han, S.J., Chung, J.H., Cheong, C.S., Chung, I.Y. and Han, Y.S. (2004) Enhancement of lipase activity from *Acinetobacter* species SY-01 by random mutagenesis and the role of lipase-specific chaperone. Enz. Microb. Technol. *35*(5), 377–384.

Henne, A., Schmitz, R.A., Bomeke, M., Gottschalk, G. and Daniel, R. (2000) Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. Appl. Environ. Microbiol. *66*, 3113–3116.

Hube, B., Stehr, F., Bossenz, M., Mazur, A., Kretschmar, M. and Schafer, W. (2000) Secreted lipases of *Candida albicans*: Cloning, characterization and expression analysis of a new gene family with at least ten members. Arch. Microbiol. *174*(5), 362–374.

Hulo, N., Sigrist, C.J.A., Le Saux, V., Langendijk-Genevaux, P.S., Bordoli, L., Gattike, A., De Castro, E., Bucher, P. and Bairoch, A. (2004) Recent improvements to the PROSITE database. Nucl. Acids. Res. *32*, D134–D137.

Jaeger, K.-E. and Eggert, T. (2002) Lipases for biotechnology. Curr. Opin. Biotechnol. *13*, 390–397.

Jaeger, K.-E. and Eggert, T. (2004) Enantioselective biocatalysis optimized by directed evolution. Curr. Opin. Biotechnol. *15*, 305–313.

Jaeger, K.-E., Dijkstra, B.W. and Reetz, M.T. (1999) Bacterial biocatalysts: Molecular biology, three dimensional structures, and biotechnological applications of lipases. Ann. Rev. Microbiol. *53*, 315–351.

Jenkins-Kruchten, A.E., Bennaars-Eiden, A., Rossi, J.R., Shen, W-J., Kraemer, F.B. and Bernlohr, D.A. (2003) Fatty acid-binding protein-Hormone-sensitive lipase interaction. Fatty acid dependence of binding. J. Biol. Chem. *278*, 47636–47643.

Kauffmann, I. and Schmidt-Dannert, C. (2001) Conversion of *Bacillus thermocatenulatus* lipase into an efficient phospholipase with increased activity towards long-chain fatty acyl substrates by directed evolution and rational design. Prot. Eng. *14*, 919–928.

Kazlauskas, R.J. (2005) Enhancing catalytic promiscuity for biocatalysis. Curr. Opinion Chem. Biol. *9*, 195–201.

Kim, H.K., Jung, Y-J., Choi, W-C., Ryu, H.S., Oh, T-K. and Lee, J-K. (2004) Sequence-based approach to finding functional lpases from microbial genome databases. FEMS Microbiol. Letters *235*, 349–355.

Kirk, O., Borchert, T.V. and Fuglsang, C.C. (2002) Industrial enzyme applications. Curr. Opin. Biotechnol. *13*, 345–351.

Klein, R.R., King, G., Moreau, R.A. and Haas, M.J. (1997) Altered acyl chain length specificity of Rhizopus delemar lipase through mutagenesis and molecular modeling. Lipids *32*(2), 123–130.

Koga, Y., Kato, K., Nakano, H., Yamane, T. (2003) Inverting enantioselectivity of *Burkholderia cepacia* KWI-56 lipase by combinatorial mutation and high-throughput screening using single-molecule PCR and *in vitro* expression. J. Mol. Biol. *331*(3), 585–592.

Kohno, M., Funatsu, J., Mikami, B., Kugimiya, W., Matsuo, T. and Morita, Y. (1996) The crystal structure of lipase II from *Rhizopus niveus* at 2.2 angstrom resolution. J. Biochem. *120*, 505–510.

Kovac, A., Scheib, H., Pleiss, J., Schmid, R.D. and Paltauf, F. (2000) Molecular basis of lipase stereoselectivity. Eur. J. Lipid Sci. Technol. *2000*: 61–77.

Lang, D.A., Mannesse, M.L.M., de Haas, G.H., Verheij, H.M. and Dijkstra B.W. (1998) Structural basis of the chiral selectivity of *Pseudomonas cepacia* lipase. Eur. J. Biochem. *254*, 333–340.

Le, G.T., Abbenante, G., Becker, B., Grathwohl, M., Halliday, J., Tometzki, G., Zuegg, J. and Meutermans, W. (2003) Molecular diversity through sugar scaffolds. Drug Discovery today: 8, 701 and refs. cited therein.

Lee, G.C., Tang, S.J., Sun, K.H. and Shaw, J.F. (1999) Analysis of the gene family encoding lipases in *Candida rugosa* by competitive reverse transcription PCR. Appl. Environ. Microbiol. *65*, 3888–3895.

Liebton, K., Zonta, A., Schimossek, K., Nardini, M., Lang, D., Dijkstra, B.W., Reetz, M.T. and Jaeger, K-E. (2000) Directed evolution of an enantioselctive lipase. Chem. Biol. *2000*(7), 709–718.

Lombardo, D., Chapus, C., Bourne, Y. and Cambillau, C. (1989) Crystallization and preliminary X-ray study of horse pancreatic lipase. J. Mol. Biol. *205*, 259–261.

Lopez, N., Pernas, M.A., Pastrana, L.M., Sanchez, A., Valero, F. and Rua, M.L. (2004) Reactivity of pure *Candida rugosa* lipase isoenzymes (Lip1, Lip2 and Lip3) in aqueous and organic media. Influence of the isoenzymatic profile on the lipase performances in organic media. Biotechnol. Prog. *20*, 65–73.

Lotti, M., Grandori, R., Fusetti, F., Longhi, S., Brocca, S., Tramontano, A. and Alberghina, L. (1993) Cloning and analysis of *Candida cylindracea* lipase sequences. Gene *124*, 45–55.

Lotti, M., Monticelli, S., Montesinos, J.L., Brocca, S., Valero, F. and Lafuente, J. (1998) Physiological control on the expression and secretion of *Candida rugosa* lipase. Chem. Phys. of Lipids *93*(1–2), 143–148.

Magnusson, A.O., Roticci-Mulder, J.C., Santagostino, A. and Hult, K. (2005) Creating space for large secondary alcohols by rational redesign of *Candida antarctica* lipase B. Chem. Bio. Chem. *6*, 1051–1056.

Mancheno, J.M., Pernas, M.A., Martinez, M.J., Ochoa, B., Rua, M.L. and Hermoso, J.A. (2003) Structural insights into the lipase/esterase behavior in the *Candida rugosa* lipases family: Crystal structure of the lipase 2 isoenzyme at 1.97 angstrom resolution. J. Mol. Biol. *332*, 1059–1069.

Martinez, C., De Geus, P., Lauwereys, M., Matthyssens, G. and Cambillau, C. (1992) *Fusarium solani* cutinase is a lipolytic enzyme with a catalytic serine accessible to solvent. Nature *356*, 615–618.

Miller, G.C., Long, C.U.J., Bojilova, E.D., Marchandier, D., Badellino, K.O., Blanchard, N., Fuki, I.V., Glick, J.M. and Rader, D.J. (2004) Role of N-linked glycosylation in the secretion and activity of endothelial lipase. J. Lipid Res. *45*(11), 2080–2087.

Mukherijee, K.D. and Hills, M.J. (1994) Lipases from plants. In: Woolley P., Petersen S.B. eds Lipases: their structure, biochemistry and application. Cambridge University Press, pp. 49–75.

Müller, G. and Petry, S. eds. Lipases and phospholipases in drug development: from biochemistry to molecular pharmacology, Wiley VCH, 2004.

Murzin, A.G., Brenner, S.E., Hubbard, T. and Chothia, C. (1995) SCOP: A structural classification of proteins databases for the investigation of sequences and structures. J. Mol. Biol. *247*, 536–540. http://scop.mrc-lmb.cam.ac.uk/scop/

Nardini, M., Lang, D.A., Liebeton, K., Jaeger, K.E. and Dijkstra, B.M. (2000) Crystal structure of *Pseudomonas aeruginosa* lipase in the open conformation - The prototype for family I.1 of bacterial lipases. J. Biol. Chem. *275*(40), 31219–31225.

Noble, M.E., Cleasby, A., Johnson, L.N., Egmond, M.R. and Frenken, L.G. (1993) The crystal structure of triacylglycerol lipase from *Pseudomonas glumae* reveals a partially redundant catalytic aspartate. FEBS Lett. *331*, 123–128.

Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G. and Goldman, A. (1992) The alpha/beta-hydrolase fold. Prot. Eng. *5*, 197–211.

Park, S. and Kazlauskas, R.J. (2003) Biocatalysis in ionic liquids-advantages beyond green technology. Curr. Opin. Biotechnol. *14*, 432–437.

Peters, G.H., Svendsen, A., Langberg, H., Vind, J., Patkar, S.A. and Kinnunen, P.K.J. (2002) Glycosylation of *Thermomyces lanuginosa* lipase enhances surface binding towards phospholipids, but does not significantly influence the catalytic activity. Colloids and Surface B-Biointerfaces *26*(1–2), 125–134.

Pleiss, J., Fisher, M. and Schmid, R.D. (1998) Anatomy of lipase binding sites : tThe scissile fatty acid binding sites. Chem. Phys. Lipids *93*, 67–80.

Qian, Z., and Lutz, S. (2005). Improving the catalytic activity of *Candida antarctica* lipase B by circular permutation. J. Am. Chem. Soc. *127* (39), 13466–13467.

Ransac, S., Rogalska, E., Gargouri, Y., Deveer, A. M., Paltauf, F., de Haas, G. H. and Verger, R. (1990). Stereoselectivity of lipases. I. Hydrolysis of enantiomeric glyceride analogues by gastric and pancreatic lipases, a kinetic study using the monomolecular film technique. J. Biol. Chem. *265*, 20263–20270.

Reetz, M.T., Bocola, M., Carballeira, J.D., Zha, D.X. and Vogel, A. (2005). Expanding the range of substrate acceptance of enzymes: Combinatorial active-site saturation test. Angew. Chem. Int. Ed. *44* (27), 4192–4196.

Reyes-Duarte, D., Polaina, J., Lopez-Cortés, N., Alcade, M., Plou, F.J., Elborough, K., Ballesteros, A., Timmis, K.N., Golyshin, P.N. and Ferrer, M. (2005) Conversion of a carboxylesterase into a triacylglycerol lipase by a random mutation. Angew. Chem. Int. Ed. *44*, 7553–7557.

Roticci, D., Roticci-Mulder, J.C., Denman, S., Norin, T., and Hult, K. (2001). Improved enantioselectivity of a lipase by rational protein engineering. Chem Biol Chem *2*, 766–770.

Roussel, A., Yang, Y.Q., Ferrato, F., Verger, R., Cambillau, C. and Lowe, M. (1998). Structure and activity of rat pancreatic lipase-related protein 2. J. Biol. Chem. *273*, 32121–32128.

Roussel, A., Canaan, Egloff, M.P., Riviere, M., Dupuis, L., Verger, R., and Cambillau, C. (1999). Crystal structure of human gastric lipase and model of lysosomal acid lipase, two lipolytic enzymes of medical interest. J. Biol. Chem. *274*, 16995–17002.

Roussel, A., Miled, N., Berti-Dupuis, L., Riviere, M., Spinelli, S., Berna, P., Gruber, V., Verger, R. and Cambillau, C. (2002). Crystal structure of the open form of dog gastric lipase in complex with a phosphonate inhibitor. J. Biol. Chem. *277*, 2266–2274.

Santarossa, G., Gatti Lafranconi, P., Alquati, C., De Gioia, L., Alberghina, L., Fantucci, P. and Lotti, M (2005). Mutations in the "lid" regions affect chain length specificity and thermostability of a *Pseudomonas fragi* lipase. FEBS Letters *579*, 2383–2386.

Sarda, L. and Desnuelle, P. (1958). Action de la lipase pancréatique sur les esters en émulsion. Biochim. Biophys. Acta *30*, 513–521.

Schmid, R. D. and Verger, R. (1998). Lipases: interfacial enzymes with attractive applications. Angew. Chem. Int. Ed. *37*, 1608–1633.

Schmidt, A., Dordick, J.S., Hauer, B., Kiener, A., Wubbolts, M. and Witholt, B. (2001). Industrial biocatalysts today and tomorrow. Nature 2001 *409*, 258–268.

Schmitt, J., Brocca, S., Schmid, R.D. and Pleiss, J. (2002). Blocking the tunnel: engineering of *Candida rugosa* lipase mutants with short chain length specificity. Prot. Eng. *15*, 595–601.

Scharg, J.D.and Cygler, M. (1993). 1.8-Angstrom refined structure of the lipase from *Geotrichum candidum*. J. Mol. Biol. *230*, 575–591.

Scharg, J.D., Li, Y., Cygler, M., Lang, D., Burgdorf, T., Hecht, H.J., Schmid, R., Schomburg, D., Rydel, T.J., Oliver, J.D., Strickland, L.C., Dunaway, C.M., Larson, S.B., Day, J. and Mcpherson, A. (1997). The open conformation of a *Pseudomonas* lipase. Structure *5*(2): 187–202.

Shibamoto, H., Matsumoto, T., Fukuda, H., and Kondo, A. (2004) Molecular engineering of *Rhizopus oryzae* lipase using a combinatorial protein library constructed on the yeast cell surface. J. Mol. Cat. B: Enzymatic (4–6), 235–239.

Shibatani, T., Nakamichi, K. and Matsumae, H. (1990). European Patent Application EP 362556.

Suen, W.C., Zhang, N.Y., Xiao, L., Madison, V. and Zaks, A. (2004) Improved activity and thermostability of *Candida antarctica* lipase B by DNA family shuffling. Prot. Eng. Des. *17*(2), 133–140.

Svendsen, A. (2000). Lipase protein engineering. Biochim. Biophys. Acta *1543*(2), 223–238.

Terzyan, S., Wang, C.S., Downs, D., Hunter, B.and Zhang, X.C. (2000). Crystal structure of the catalytic domain of human bile salt activated lipase. Prot. Sci. *9*, 1783–1790.

Tyndall, J.D.A., Sinchaikul, S., Fothergill-Gilmore, L.A., Taylor, P., and Walkinshaw, M.D. (2002). Crystal structure of a thermostable lipase from *Bacillus stearothermophilus* P1. J. Mol. Biol. *323*, 859–869.

Uppenberg, J., Patkar, S., Bergfors, T. and Jones, T.A. (1994). Crystallization and preliminary X-ray studies of lipase B from *Candida antarctica*. J. Mol. Biol. *235*, 790–792.

van Kampen, M.D. and Egmond, M.R. (2000). Directed evolution: from a staphylococcal lipase to a phospholipase. Eur. J. Lipid Sci. Technol. *102*, 717–726.

van Pouderoyen, G., Eggert, T., Jaeger, K.-E. and Dijkstra, B.W. (2001). The crystal structure of*Bacillus subtilis* lipase: a minimal alpha/beta hydrolase fold enzyme. J. Mol. Biol. *309*(1), 215–226.

van Tilbeurgh, H., Sarda, L., Verger, R. and Cambillau, C. (1992). Structure of the pancreatic lipase–procolipase complex Nature *359*, 159–162.

Verger, R. (1997). "Interfacial activation" of lipases: facts and artifacts. Trends Biotechnol. *15*, 32–38.

Villeneuve, P., Muderhwa, J.M., Graille, J. and Haas, M.J. (2000). Customizing lipases for biocatalysis: a survey of chemical, physical and molecular biological approaches. J. Mol. Catal. B: Enzymatic *9*, 113–148.

Voget, S., Leggewie, C., Uesbeck, A., Raasch, C., Jaeger, K.-E. and Streit, W.R. (2003). Prospecting for novel biocatalysts in a soil metagenome. Appl. Environ. Microbiol. *69*, 6235–6242.

Wei, Y., Swenson, L., Castro, C., Derewenda, U., Minor,W., Arai, H., Aoki, J., Inoue, K., Servin-Gonzalez, L. and Derewenda, Z.S. (1998). Structure of a microbial homologue of mammalian platelet-activating factor acetylhydrolases: *Streptomyces exfoliatus* lipase at 1.9 Å resolution. Structure *6*, 511–519.

Wicker-Planquart, C., Canaan, S., Rivière, M. and Depuis, L. (1999). Site-directed removal of N-glycosylation sites in human gastric lipase. Eur. J. Biochem. *262*, 644–651.

Winkler, FK, D'Arcy, A. and Hunziker W. (1990). Structure of human pancreatic lipase. Nature *343*, 771–774.

Yeaman, S. J. (2004). Hormone-sensitive lipase – new roles for an old enzyme. Biochem. J. *379*, 11–22.

Zschenker, O., Bahr, C., Hess, U.F. and Ameis, D. (2005). Systematic mutagenesis of potential glycosylation sites of lysosomal acid lipase. J. Biochem. *137*(3), 387–394.

Zaks, A., and Klibanow, A. M. (1984). Enzymatic catalysis in organic media at 100 °C. Science *224*, 1249–1251.

Zha, D.X., Wilensek, S., Hermes, M., Jaeger, K.-E. and Reetz M.T. (2001). Complete reversal of enantioselectivity of an enzyme-catalyzed reaction by directed evolution. Chem. Commun. *2001*, 2664–2665.

CHAPTER 17

USE OF LIPASES IN THE INDUSTRIAL PRODUCTION OF ESTERS

SOUNDAR DIVAKAR* AND BALARAMAN MANOHAR
Central Food Technological Research Institute, Mysore, India
**divakar643@gmail.com*

1. INTRODUCTION

Of the estimated 25,000 enzymes present in nature only about 2800 have been characterized, and of these about 400, mainly hydrolases, transferases and oxidoreductases, have been identified as being potentially useful from a commercial point of view. However, only 50 different kinds of enzymes are employed on an industrial scale (Berger, 1995). Lipases (triacyl glycerol hydrolases E.C.3.1.1.3) catalyse the hydrolysis of triglycerides at the oil/water interface, but their ability to form ester bonds under reverse hydrolytic conditions enables them to catalyse various other types of reactions such as esterification, transesterification, polymerisation and lactonization. The high selectivity and mild conditions associated with lipase-mediated transformations have made them very attractive for the synthesis of a wide range of natural products, pharmaceuticals, fine chemicals, food ingredients (Schreier, 1997) and bio-lubricants (Dörmö *et al*, 2004). For example, lipases are employed to obtain polyunsaturated fatty acids (PUFAs) which are then used along with mono- and diglycerides for the synthesis of nutraceuticals and pharmaceuticals such as anticholesterolemics, anti-inflammatories and thrombolytics (Gill and Valivety, 1997; Belarbi *et al*, 2000). The main reason for the use of lipases is the growing interest and demand for products prepared by natural and environmentally compatible means. As a consequence of their versatility in application, lipases are regarded as enzymes of high commercial potential. Lipase catalysed esterification in organic solvents presents challenges, which if dealt with successfully, can result in the generation of a number of useful compounds.

Both the range of substrates with which lipases react and the range of reactions catalysed are probably far wider than those of any other enzyme studied to date.

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 283–300

Table 1. Lists of some commercially important esters produced by lipases

Name of compound	Use	Lipase	References
A. Flavour Esters			
Isoamyl acetate	Banana flavour	*Candida antarctica* *Rhizomucor miehei* *Aspergillus niger* *Pseudomonas pseudomallei* Lipolase 100T, Novozym 435	Langrand *et al.,* 1990 Chulalaksananukul *et al.,* 1993 Welsh *et al.,* 1990 Kanwar and Goswami 2002 Kumar *et al.,* 2005
Isoamyl butyrate	Banana flavour	*Candida antarctica* *Rhizomucor miehei* *Candida cylindraceae*, PPL, *Aspergillus niger*	Langrand *et al.,* 1990 Mestri and Pai, 1994b
Isoamyl propionate	Banana flavour	*Rhizomucor miehei*	Chowdary *et al.,* 2002
Isoamyl isovalerate	Apple flavour	*Rhizomucor miehei*	Chowdary *et al.,* 2002
Isobutyl isobutyrate	Pineapple flavour	*Rhizomucor miehei*	Hamsaveni *et al.,* 2001
Methyl propionate	Fruity flavour	*Rhizomucor miehei*	Perraud and Laboret 1989
Ethyl butyrate	Pineapple flavour	*Candida cylindracea*	Yadav and Lathi 2003
Butyl isobutyrate	Sweet fruity odour	*Candida cylindracea*, PPL and *Aspergillus niger*	Welsh and Williams 1990
Protocatechuic aldehyde		*Rhizomucor miehei*, PPL	Divakar, 2003
Short chain fatty acid esters	Fruity odour		Mestri and Pai 1994a Xu *et al.,* 2002
Long chain alcoholic esters of lactic acids	Flavour	*Candida antarctica*	From *et al.,* 1997; Torres and Otero, 1999
Methyl benzoate	Exotic fruity and berry flavour	*Candida rugosa*	Leszczak and Tran-Minh 1998
Tetrahydrofurfuryl butyrate	Fruity favour	Novozym 435	Yadav and Devi, 2004
Cis-3-hexen-1-yl acetate	Fruity odour	*Rhizomucor miehei*	Chiang *et al.,* 2003

B. Fragrance Esters			
Tolyl esters	Honey note	*Rhizomucor miehei*, PPL	Suresh Babu *et al.,* 2002; Manohar and Divakar, 2002,
Anthranilic acid esters of C_2–C_{18} alcohols	Flowery odour of jasmine	*Candida cylindracea* PPL	Kittleson and Pantaleone, 1994 Suresh Babu and Divakar, 2001
4-*t*-Butylcyclohexyl acetate	Woody and intense flowery notes	PPL	Manohar and Divakar, 2004b
Geranyl methacrylate	Floral fruity odour	*Rhizomucor miehei*, PPL, *Pseudomonas cepacia*	Athawale *et al.,* 2002
Citronellyl acetate	Fruity rose odour	*Candida antarctica* SP435	Claon and Akoh, 1994
Citronellyl propionate		*Pseudomonas fragi*	Mishio *et al.,* 1987
Citronellyl valerate			Marlot *et al.,* 1985
Farnesol and phytol esters	Fruity odour	*Candida rugosa*	Shieh *et al.,* 1996
α-Terpinyl esters	Fruity, characteristic	*Rhizomucor miehei*	Rao and Divakar, 2002
α-Terpinyl acetate	lavendar and	*Rhizomucor miehei*	Rao and Divakar, 2001
α-Terpinyl propionate	bergamot-like fragrance	*Aspergillus niger, Rhizopus delemar, Geotrichum candidum, Pencillium cyclopium*	Claon and Akoh, 1994
α-Terpinyl esters of fatty acids			
α-Tetpinyl esters of short chain acids			
Terpinyl esters of triglycerols			
C. Surfactant Esters			
Oleic acid esters of terpinyl alcohols	Surfactants		Okumura *et al.,* 1979
Oleic acid esters of short chain alcohols	Surfactants	Novozym 435	Dörmö *et al.,* 2004
Butyl oleate	Surfactants	*Rhizomucor miehei*	Knez *et al.,* 1990
2-O- Alkanoyl lactic acid esters of C_2–C_{18} alcohols	Surfactants	*Rhizomucor miehei*, PPL	Kiran and Divakar, 2001

(Continued)

Table 1. (Continued)

Name of compound	Use	Lipase	References
D. Surfactant and Sweeteners			
N-Acetyl- L-leucyl, L-methionyl, L-tyrosinyl, L-tryptophnyl–D–glucose	Surfactants	*Mucor javanicus*, *Pseudomonas cepacia*, Subtilisin,	Maruyama *et al.,* 2002
N-Acetyl-L-phenylalanyl esters of carbohydrates		Subtilisin	Maruyama *et al.,* 2002; Riva *et al.,* 1988
N-Acetyl-L-methionyl-methyl-β-galactopyranoside		Optimase M-440, Proleather, APG 380	Park *et al.,* 1996
N-t-Boc-L-phenylalanyl esters of carbohydrates		Optimase M-440	Park *et al.,* 1999
N-t-Boc-L-leucyl, L-tyrosinyl, -L-methionyl, L-aspartyl, L-lysyl-sucrose			
N-t-Boc-L-phenylalanyl esters of sugar alcohols		Optimase M-440	Jeon *et al.,* 2001
L-Alanyl, L-Leucyl esters of D-glucose	Surfactants	*Rhizomucor miehei*, PPL	Vijayakumar *et al.,* 2004
L-Phenylalanyl esters of D-glucose	Surfactants	*Rhizomucor miehei*, PPL	Vijayakumar *et al.,* 2004; Lohith, 2005
N-Acetyl-L-alanyl esters of carbohydrates	Surfactants and sweeteners	Subtilisin *Rhodotorula lactosa* PPL Subtilisin	Riva *et al.,* 1988 Suzuki *et al.,* 1991 Klibanov, 1986; Boyer *et al.,* 2001
Fructose oleate Fatty acid esters of glycosides	Surfactant	Lipozyme, *Rhizomucor miehei*, *Candida antarctica*	Khaled *et al.,* 1991 Adlerhorst *et al.,* 1990
Butyl oleate Oleyl butyrate Oley oleate	Surfactants	Candida rugosa	Zaidi *et al.,* 2002
Lauroyl esters of carbohydrates Palmitoyl esters of maltose	Surfactants	*Humicola lanuginose* *Candida antarctica* B	Ferrer *et al.,* 2005

β-Methylglucoside methacrylate/acrylate	Surfactants	*Candida antarctica*	Kim *et al.,* 2004
E. Bio-degradable Polyesters	Molecular weight		
Ring opening polymerization of α-methyl-β-propiolactone	600–2900	*Pseudomonas fluorescens*	Svirkin *et al.,* 1996
Copolymerization of β-propiolactone and ε-caprolactone	520	*Pseudomonas fluorescens*	Namekawa *et al.,* 1996
Poly-ε-caprolactone ester	7600	PPL	Henderson *et al.,* 1996
Polymerization of Macrolides:			
Octanolide			Kobayashi *et al.,* 1998
Undecanolide	25000		Uyama *et al.,* 1995
Dodecanolide			Uyama *et al.,* 1995
Pentadecanolide			Uyama and Kobayashi, 1996; Bisht *et al.,* 1997
Hexadecanolide			Namekawa *et al.,* 1996
Polymerization of lactic acid	1423	PPL	Kiran and Divakar, 2003
Poly-ε-caprolactone	11000	PPL	Divakar, 2004

Lipases catalyse three types of reactions: i) **Hydrolysis**: occurs in aqueous media when there is large excess of water, ester hydrolysis is the dominant reaction; ii) **Esterification**: under low water conditions such as in nearly anhydrous solvents, esterification can be achieved (improved product yields can be obtained if the water content of the medium is controlled); iii) **Transesterification**: the acid moiety of an ester is exchanged with another one (if the acyl donor is a free acid the reaction is called acidolysis, whereas the reaction is called interesterification if the acyl donor is an ester; in alcoholysis, the nucleophile alcohol acts as an acyl acceptor). Lipases are currently used in the production of many commercially important esters. Table 1 shows a list of these products and includes flavours, fragrances, surfactants, sweeteners and biodegradable polyesters.

2. FACTORS AFFECTING LIPASE-MEDIATED ESTERIFICATION

2.1. Nature of Substrate

Lipases display varying degrees of selectivity towards the substrates with which they interact. Steric hindrance (branching, unsaturation and chain length) and electronic effects of the substrates are the two major factors that determine selectivity. In esterification reactions, many lipases display high selectivity for long and medium chain fatty acids rather than short chain or branched ones (Alhir *et al*., 1990). Most lipases display selectivity towards carboxylic acids. *G. candidum* lipase reacts only with fatty acids containing a *cis* bond at the 9th position (Schrag *et al*., 1996). Alcohols like ethanol and geraniol have been reported to be inhibitory in esterification and transesterification reactions. Substrate molar ratio plays an important role in the esterification reaction. The latter can be improved by increasing the amount of either alcohol or acid present but in most cases alcohols may be inhibitory and acids may cause acidification of the microaqueous interface resulting in inactivation of lipases (Dörmö *et al*., 2004; Zaidi *et al*., 2002). It is difficult to generalize the effect of chain length on esterification because this depends on individual lipase preparations and the specificity of the enzymes. Esterification increased with increasing chain length in reactions catalysed with lipases from *Staphylococcus warneri* and *Staphylococcus xylosus*. In the case of Lipolase 100T esterification decreased with increasing chain length, and was found to be independent of chain length when catalysed with Novozyme 435 (Kumar *et al*., 2005).

The use of acetic acid as an acyl donor in the preparation of acetates was attempted with little or no success. Compared to longer chain carboxylic acids (propionates, butyrates), acetic acid is a potent lipase inhibitor (Segel, 1975), preferentially reacting with the serine residue at the active site (Huang *et al*., 1998). It was not possible to observe any reaction between acetic acid and geraniol using lipases from different micro-organisms. It was also shown that acetic acid esters were difficult to synthesize at high yield due to lipase inactivation by acid. While some researchers have focused their attention on transesterification to obtain high yields of acetates (Chulalaksananukul *et al*., 1993), reports on maximizing acetate

production by direct esterification are scant. Also, low molecular weight substrates are more water-soluble and as such may react differently to high molecular weight (less water soluble) substrates in non-aqueous systems.

2.2. Nature of Solvent

Most information regarding enzyme catalysis such as reaction rates, kinetics and mechanistic aspects have been derived from studies conducted in aqueous solutions (Welsh and Williams, 1990). However, when enzymes are directly dispersed in organic solvents they exhibit remarkable changes in their properties (Klibanov, 1986). Organic solvents influence reaction rate, maximum velocity (V_{max}), specific activity (K_{cat}), substrate affinity (K_M), specificity constants (K_{cat}/K_M) enantio-selectivity, lipase stability and stereo- and regio-selectivities (Sakurai *et al*., 1988; Kung and Rhee, 1989; Zaks and Klibanov, 1986). Differences in enzyme activity in different solvents could be due to variable degrees of enzyme hydration imposed by the solvents rather than a direct effect on the enzyme or substrates.

Studies on the quantification of solvent effects on enzyme catalysis have been carried out (Laane *et al*., 1987). Employing the Hildebrand parameter, δ, as a measure of solvent polarity it was concluded originally that enhanced reaction rates could be expected when the polarity of the organic solvent was low ($\delta \approx 8$) and molecular weight grater than 150. Nevertheless, δ was later demonstrated to be a poor measure of solvent polarity. Laane *et al*. (1987) quantified solvent polarity on the basis of log P values. The log P value of a solvent is defined as the logarithm of the partition coefficient of the solvent in an *n*-octanol/water two-phase system. Generally, biocatalysis is low in solvents of $\log P < 2$, is moderate in solvents with a log P value between 2 and 4 and high in non-polar solvents of $\log P > 4$. *Rhizomucor miehei* lipase was shown to conform to these rules when esterification reactions were conducted in different solvents (Laane *et al*., 1987). In the presence of hydrophilic solvents ($\log P < 2$) lipozyme showed no esterification. Hence, polar solvents may remove the essential water from the enzyme and disrupt the active confirmation (Adachi and Kobayashi, 2005). Solvents of $\log P > 2$ dissolve to a lesser degree in water, leaving the enzyme suitably hydrated in its active conformation and hence are able to support product synthesis (Soo *et al*., 2003). A lipase which exhibits increasing activity with increased content of DMSO - a polar solvent - has also been isolated. The influence of log P of organic solvents was studied by correlating these with K_{cat} and K_m values. K_{cat} showed strong correlation with log P whereas K_m did not. K_{cat} was not affected by different solvent compositions having the same log P value whereas K_m was reported to change remarkably (Hirakawa *et al*., 2005). Polar solvents besides inactivating lipases dissolve certain alcohols like sugars, therefore mixtures of non-polar solvents containing a small amount of polar solvents have been employed in lipase catalysis involving sugars. While it is generally accepted that non-polar solvents are better than polar ones for lipase catalysed esterification

reactions, a clear consensus has yet to be reached regarding the issue of solvent effects on enzyme catalysis in general.

2.3. Thermal Stability

Many factors govern the catalytic activity and operational stability of lipases at higher temperatures in non-aqueous media. Two of the most important concern the nature of the organic medium employed and the water content in the microenvironment of the enzyme. There are a few reports on the thermostability of lipases in aqueous media. The lipase from *Pseudomonas fluorescens* 33 was found to retain 10–20% more activity during heating to 60 °C–90 °C for 10 min when casein and Ca^{2+} were present (Kumura *et al*., 1993). The thermostabilities of some serine esterases such as chymotrypsin and lipase from *Candida rugosa* and *Rhizomucor miehei* have been studied as a function of enzyme hydration using differential scanning calorimetry (Turner *et al*., 1995). It was found that the denaturation temperature (T_m) was 30 °C–50 °C higher in anhydrous environments compared to aqueous solutions. Porcine pancreas lipase was reported to retain greater esterification activity in a dry organic environment (2M heptanol solution in tributyrin) at a temperature of 100 °C when a low concentration of water (0.015%) was maintained in the reaction system. The half-life of the enzyme was found to be more than 12 h at 100 °C. However, when the concentration of water was increased to 3%, loss of activity was almost instantaneous (half life = 2 min). Porcine pancreas lipase (PPL) in non-aqueous media showed that long periods of incubation (up to 10 days) at 80 °C did not affect the active conformation of PPL (Kiran *et al*., 2001a). Immobilization and the addition of salt hydrates are known to enhance the thermostability of lipases in organic media. Thermal stability can also be improved by making surfactant-lipase complexes (Goto *et al*., 2005). It is common practice now to carry out lipase catalysed esterification reactions at around 80 °C–90 °C. Noel and Combes (2003) conducted a series of experiments to study the effects of temperature on *Rhizomucor miehei* lipase (RML) and concluded that thermal deactivation occurs due to the formation of aggregates rather than protein unfolding.

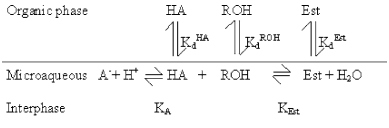
2.4. Role of Water in Lipase-mediated Catalysis

Water plays a crucial role in the reversible reaction catalysed by lipase (Gayot *et al*., 2003). While a critical amount of water is necessary for maintaining the active conformation of the enzyme, excess water facilitates hydrolysis (Cameron *et al*., 2002). Bound water is very important in stabilizing the conformation of a lipase in non-aqueous media. In the case of the *Rhizomucor miehei* lipase, water is bound to charged and polar residues on the surface of the enzyme as a monolayer (Tramper *et al*., 1992). The presence of excess water decreases the catalytic activity from both the kinetic and thermodynamic points of view. The concentration of water in organic solvents is inversely proportional to the thermostability of lipases. It was shown that for PPL, hydrophobic solvents served better than hydrophilic ones for catalysis.

Substrate concentrations and water activity can determine product distribution, hence the monoester of ethylene glycol can be prepared by using either low water activity or by employing higher concentrations of alcohols, and *vice versa* for diester synthesis (Chand *et al*., 1997). Osorio *et al*. (2001) reported that beyond a critical water concentration, lipase-mediated esterification decreases because the extent of the water layer formed around the enzyme retards the transfer of the acyl donor to the active site of the enzyme. Yadav and Devi (2004) conducting experiments at various agitation speeds found that there is no effect of the speed of agitation on esterification. The water layer surrounding the enzyme makes the latter more flexible, acting as a molecular lubricant by forming multiple hydrogen bonds with it. However, beyond a certain critical level, increased amounts of water may result in excessive flexibility resulting in interaction between the enzyme and the organic solvent with consequent denaturation of the former and loss of activity. In addition, organic substrates and products which are poorly soluble in aqueous media diffuse with difficulty through the intra-particle water layer to the active centre of the enzyme. Thus the activity of the enzyme would be influenced by both water-induced inactivation and partition of components between the bulk solvent and the microenvironment of the lipase (Yadav and Devi, 2004). Almost all lipases are active at low water activity but there are large differences in optimal water activity between them (Ma *et al*., 2002).

Several methods are available to monitor water activity such as Karl-Fischer titration and specialized sensors. In esterification reactions, the water formed can be removed by passing the reaction mixture through a bed of desiccants leading to greater product yields. In a non-polar solvent, excess water adds to the already existing hydration shell on the enzyme constituting the microaqueous interface. Partitioning of the acid, alcohol and product between the microaqueous interface and the solvent phase plays a significant role in regulating esterification. The solubility of the acid and its dissociation results in a build-up of protons at the interface. In lipase-catalysed esterification, the various equilibria involved at the microaqueous interface are shown in Scheme 1 (Aires-Barros *et al*., 1989). Where HA = acid, ROH = alcohol, Est = ester, K_d^{HA}, K_d^{ROH} and K_d^{Est} = distribution coefficients of acid, alcohol and ester respectively; K_A = dissociation constant of acid, K_{Est} = equilibrium constant of esterification.

Since water is present in micro-quantities and is inaccessible, direct measurement of microaqueous pH is not possible (Valivety *et al*., 1990). Attempts have been made to measure the pH in non-aqueous systems by Cambou and Klibanov (1984) who reported the use of an indicator which changed colour with pH. A reliable



Scheme 1. Equilibria operating at the microaqueous interphase in the lipase catalysed esterification in organic solvents

method has been developed by Valivety *et al.* (1990) using a hydrophobic indicator (fluorescein ester with 3,7,11- trimethyldodecanol) which remains completely in the organic phase but responds to pH changes in an adjacent aqueous phase. Thermodynamic factors operating at the enzyme-water-solvent interface in non-polar solvents have also been investigated in terms of the water of reaction, partitioning of acid between the microaqueous phase and the organic solvent, dissolution and dissociation of the acid, the resultant number of H^+ present in the microaqueous phase and the extent of esterification (Kiran *et al.*, 2002).

3. KINETIC STUDIES OF LIPASE CATALYSED ESTERIFICATION REACTIONS

The kinetics of lipase-catalysed esterification reactions help in not only quantifying a reaction but also reveal details of enzyme inhibition and mechanism which have quite a lot of bearing on suitability in industrial applications. Lipases used in organic solvents followed a complex two-substrate Ping-Pong Bi-Bi mechanism. A Ping-Pong Bi-Bi mechanism, which stands for two-substrate two-product reaction, is a sequential one *i.e.* both substrates do not bind to the enzyme simultaneously before the product is formed (Segel, 1975). The amount of lipase available and the rate of breakdown of the enzyme-substrate complex govern the overall rate of reaction. If the organic acid employed is inhibitory in nature then it remains bound to the enzyme strongly and no acyl transfer occurs. In some cases, even if acyl transfer occurs, the product formed may remain bound to the enzyme resulting in inhibition.

Lipase-catalysed esterification between oleic acid and ethanol and transesterification between geraniol and propyl acetate (Chulalaksananuku *et al.*, 1992) were found to follow a Ping-Pong Bi-Bi mechanism where both ethanol and geraniol were found to be inhibitory. A similar Ping-Pong Bi-Bi mechanism was found to be followed in the kinetics of esterification of lauric acid by (−)-menthol catalysed by the lipase from*Penicillium simplicissium*, with (−)-menthol being inhibitory (Stamatis *et al.*, 1993). In a transesterification reaction between isoamyl alcohol and ethyl acetate catalysed by Lipozyme IM20, the substrates ethyl acetate and isoamyl alcohol and one of the products (ethanol) were found to be inhibitory. Of the three, ethanol was found to be the greatest inhibitor. Thus, improved kinetic models being proposed will allow to predict enzyme behaviour.

4. ENZYME IMMOBILIZATION

Enzyme immobilization increases the number of enzyme molecules per unit area increasing the efficiency of the reaction. Like with other enzymes, the advantages of immobilizing lipases include the repetitive use of a given batch of enzyme, better process control, enhanced stability, enzyme-free products (Rahman *et al.*, 2005), increased stability of polar substrates, shifting of thermodynamic equilibria to favour ester synthesis over hydrolysis, reduction of water dependent side reactions such as hydrolysis, elimination of microbial contamination and the potential for use directly

within a chemical process. In the presence of organic solvents, immobilized lipase showed enhanced activity (Ye *et al*., 2005).

The immobilization of lipases has been performed by various methods such as adsorption, entrapment and covalent binding, using different supports. For covalent immobilization, support matrices such as silica beads are usually activated with glutaraldehyde (Ulbrich *et al*., 1991). In the case of non-covalent immobilization, lipases can be adsorbed onto a weak anion exchange resin maintaining very good activity (Ison *et al*., 1990). For non-covalent immobilization, both ionic and hydrophobic interactions between the lipase and the support surface are important. Polymers such as polyvinyl alcohol (PVA), carboxymethyl cellulose (CMC), poly ethylene oxide (PEO) and CMC/PVA blends can also be used for lipase immobilization (Dalla-Vecchia *et al*., 2005). The morphology of film surfaces analysed by scanning electron microscopy indicated that lipases were preferentially located on the polymer surface (Crespo *et al*., 2005). Dalla-Vecchia *et al*. (2005) have immobilized 10 different lipases on polyvinyl alcohol, carboxymethyl cellulose and PVA/CMC blend (50:50% m/m), and among them *Mucor javanicus* lipase (MJL) and *Rhizopus oryzae* lipase (ROL) exhibited the highest activities. Immobilized enzymes can be reused many times: *Candida antarctica* B (Novozym 435) was immobilized on mesoporous silica with octyltriethoxysilane and it retained its activity even after 15 reaction cycles (Blanco *et al*., 2004). Calcium carbonate was found to be the most suitable adsorbent when crude *Rhizopus oryzae* lipase was immobilized on different supports and it exhibited long-chain fatty acid specificity (Ghamgui *et al*., 2004). The lipase from *Pseudomonas cepacia* was gel-entrapped by polycondensation of hydrolysed tetramethoxy silane and iso-butyltrimethoxy silane and was subjected to repeated use without loosing much of its activity (Noureddini *et al*., 2005).

5. ESTERIFICATION IN REVERSE MICELLES

Enzymatic reactions in reverse micelles (water-in-oil) offer many advantages over those in micelles (oil-in-water) or in organic solvents, such as the solubilization of lipases and both hydrophobic/hydrophilic substrates at higher concentrations, better control over water activity, and a large interfacial area leading to enhanced reaction rates in a thermodynamically stable single phase (Stamatis *et al*., 1999). Various reactions including the syntheses of flavour esters and macrocylic lactones, and the resolution of chiral alcohols (Rees and Robinson, 1995) have been attempted in reverse micelles. Krieger *et al*. (2004) highlighted some of the recent developments on the use of lipases in reverse micelles. Some efforts have been made towards achieving continuous product recovery and also enzyme reuse, both of which are major problems with enzyme catalysis in reverse micelles. Reverse micelles can exchange biocatalyst, water, substrates and products with the bulk organic solvent (Krieger *et al*., 2004). The effective diffusion coefficient of lauric acid varied depending on the composition of the lecithin microemulsion-based organogels (MBGs), while that of butyl alcohol remained constant in the esterification of lauric

acid with butyl alcohol catalysed by *Candida rugosa* lipase (Nagayama *et al*, 2002). A high initial reaction rate was obtained under extremely low water content conditions when the esterification of oleic acid with octyl alcohol catalysed by *Rhizopus delemar* lipase was carried out in a reverse micelle system of sugar ester DK-F-110 (Naoe *et al*, 2001). Kinetic studies were carried out to examine the esterification of octanoic acid with 1-octanol catalysed by *Candida lypolytica* (CL) lipase, in a water-in-oil microemulsions formed by water/bis-(2-ethylhexyl) sulphosuccinate sodium (AOT)/isooctane (Zhou *et al*, 2001). An esterification reaction of hexanol and hexanoic acid in a cyclohexane/dodecylbenzenesulphonic acid (DBSA)/water microemulsion system using *Candida cylindracea* lipase demonstrated that DBSA itself can act as a kind of acid catalyst (Han and Chu, 2005).

6. RESPONSE SURFACE METHODOLOGY (RSM) IN LIPASE CATALYSIS

Response surface methodology analysis provides an important tool for parameter optimisation and has been applied to several esterification reactions. RSM studies have centred on working out the optimal conditions for particular lipase-catalysed esterification reactions. Thus, optimum conditions for the enzymatic synthesis of geranyl butyrate using lipase AY from *Candida rugosa* were worked out by Sheih *et al.* (1996). Similarly, the effect of reaction parameters on SP 435 lipase-catalysed synthesis of citronellyl acetate in organic solvents was carried out by Claon and Akoh (1994), and the optimisation of conditions for the synthesis of 2-O-palmitoyl lactic acid, 2-O-stearoyl lactic acid and 2-*O*-lauroyl lactic acid using lipases from *Rhizomucor miehei* and porcine pancreas was studied by Kiran *et al.* (2001b). RSM has also been employed to optimise the lipase-catalysed synthesis of flavours (Nogales *et al*, 2005), biodiesel (Shieh *et al*, 2003; Chang *et al*, 2005), and propylene glycol monolaurate (Shaw *et al*, 2003). The usefulness of several statistical methods including Box-Behnken, Central Composite Rotatable and Plackett-Burman designs have also been exploited for the experimental optimization of lipase catalysed esterification reactions (Manohar and Divakar, 2004a).

7. LIPASE CATALYSED RESOLUTION OF RACEMIC ESTERS

Lipases have been extensively used in the resolution of racemic alcohols and carboxylic acids through asymmetric hydrolysis of the corresponding esters. Chirally pure hydroxyalkanoic acids which find wide applications as drug intermediates have been obtained from racemic (±)-hydroxyalkanoic esters (Engel *et al.*, 1991). Molecular modelling studies have revealed that enzyme behaviour towards racemic substrates can be predicted. Rantwijk (2004) critically reviewed the resolution of chiral amines by enantioselective acylation by three different serine hydrolases such as lipases, subtilisin and *Penicillin* acylase and recommended *Candida antarctica* lipase because of its high enantioselectivity and

stability. Resolution of some enantiomeric alcohols like (*R,S*)-2-octanol, (*R,S*)-2-(4-chlorophenoxy) propionic and (*R,S*)-2-bromo hexanoic acids was carried out using lipases from *Candida rugosa* and *Pseudomonas* sp., where *R*-alcohol was obtained with an enantiometric excess of about 98% (Crespo *et al*., 2005). Optically active (S)-α-cyano-3-phenoxybenzyl (CPB) acetate was obtained from racemic cyanohydrins by transesterification using the lipase from *Alcaligenes* sp. in organic media (Zhang *et al*., 2005). A lipase-like enzyme isolated from porcine pancreas immobilized in DEAE-Sepharose gave pure (S)-(−)glycidol from (R)-(−)-glycidyl butyrate when the reaction was carried out at pH 7.0, in 10% dioxane at 25 °C (Palomo *et al*., 2003).

8. CONCLUDING REMARKS

Lipases constitute some of the most thoroughly studied hydrolysing enzymes in synthetic reactions, and have come a long way in establishing themselves as an important synthetic tool for bio-organic researchers. Whilst the synthetic applications of these enzymes in the preparation of oils, fats, structured health lipids, pharmaceuticals and other such esters are many, this chapter has of necessity focussed on just a limited set of selected ester preparations using lipases.

ACKNOWLEDGEMENT

The authors acknowledge Mr. K. Lohith for his assistance in the preparation of this manuscript.

REFERENCES

Adachi, S. and Kobayashi, T. (2005). Synthesis of esters by immobilized-lipase-catalyzed condensation reaction of sugars and fatty acids in water-miscible organic solvent. J. Biosci. Bioeng. *99*(2), 87–94.

Adlerhorst, K., Björking, F., Godtfredsen, S.E. and Kirk, O. (1990). Enzyme catalyzed preparation of 6-O-acylglucopyranosides. Synthesis *1*, 112–115.

Aires-Barros, M.R., Cabral, J.M.S., Willson, R.C., Hamel, J. F.P. and Cooney, C.L. (1989). Esterification-coupled extraction of organic acids. Partition enhancement and underlying reaction and distribution equilibria. Biotechnol. Bioeng. *34*, 909.

Alhir, S., Markajis, S. and Chandan, R. (1990). Lipase of *Penicillium caseicolum*. J. Agric. Food Chem. *38*, 598–601.

Athawale, V., Manjrekar, N. and Athawale, M. (2002). Lipase-catalyzed synthesis of geranyl methacrylate by transesterification: study of reaction parameters. Tetrahedron Lett. *43*, 4797–4800.

Belarbi, E.H., Molina, E. and Chisti, Y. (2000). A process for high yield and scaleable recovery of high purity eicosapentaenoic acid esters from microalgae and fish oil. Enzyme Microb. Technol. *26*, 516–529.

Berger, R.G. (1995). Aroma Biotechnology. Springer-Verlag. Berlin Heidelberg. NY, USA.

Bisht, K.S., Deng, F., Gross, R.A., Kaplan, D.L. and Swift, G. (1997). Ethyl glucoside as a multifunctional initiator for enzyme-catalyzed regio-selective lactone ring-opening polymerization. J. Am. Chem. Soc. *120*, 1363–1367.

Blanco, R.M., Terreros, P., Fernandez-Perez, M., Otero, C. and Díaz-Gonalez, G. (2004). Functionalization of mesoporous silica for lipase immobilization Characterization of the support and the catalysts. J. Mol. Catal., B Enzym. *30*, 83–93.

Boyer, V., Stanchev, M., Fairbanks, A.J. and Davis, B.G. (2001). Ready protease catalysed synthesis of carbohydrate-amino acid conjugates. Chem. Commun. 1908–1909.

Cambou, B. and Klibanov, A.M. (1984). Lipase-catalyzed production of optically active acids via asymmetric hydrolysis of esters. Effect of the alcohol moiety. Appl. Biochem. Biotechnol. *9*, 255–260.

Cameron, P.A., Davison, B.H., Frymier, P.D. and Barton, J.W. (2002). Direct transesterification of gases by dry immobilized lipase. Biotechnol. Bioeng. *78*, 251–256.

Chand, S., Adlercreutz, P. and Mattiasson, B. (1997). Lipase-catalysed esterification of ethylelene glycol to mono and diesters. The effect of process parameters on reaction rate and production. Enzyme Microb. Technol. *20*, 102–106.

Chang, H.M., Liao, H.F., Lee, C.C. and Sheih, C.J. (2005). Optimized synthesis of lipase-catalyzed biodiesel by Novozym 435. J. Chem. Technol. Biotechnol. *80*, 307–312.

Chiang, W.D., Chang, S.W. and Shieh, C.J. (2003). Studies on the optimized lipase-catalyzed biosynthesis of cis-3-hexen-1-yl acetate in n-hexane. Process Biochem. *38*, 1193–1199.

Chowdary, G.V., Divakar. S. and Prafulla S.G. (2002). Modeling on isoamyl isovalerate synthesis from *Rhizomucor miehei* lipase in organic media: optimization studies. World J. Microbiol. Biotechnol. *18*, 179–185.

Chulalaksanaukul, W., Condort, J.S. and Combes, D. (1992). Kinetics of geranyl acetate synthesis by lipase catalyzed transesterification in *n*-hexane. Enzyme Microb. Technol. *14*, 293–298.

Chulalaksananukul, W., Condoret, J.S. and Combes, D. (1993). Geranyl acetate synthesis by lipase catalysed transesterification in supercritical carbon dioxide. Enzyme Microb. Technol. *15*, 691–698.

Claon, P.A. and Akoh, C.C. (1994). Effect of reaction parameters on SP435 lipase-catalyzed synthesis of citronellyl acetate in organic solvent. Enzyme Microb. Technol. *16*, 835–838.

Crespo, J.S., Queiroz, N., Nascimento, M.G. and Soldi, V. (2005). The use of lipases immobilized on poly (ethylene oxide) for the preparation of alkyl esters. Process Biochem. *40*, 401–409.

Dalla-Vecchia, R., Sebrão, D., Nascimento, M.G. and Soldi, V. (2005). Carboxymethylcellulose and poly (vinyl alcohol) used as a film support for lipases immobilization. Process Biochem. *40*, 2677–2682.

Divakar, S. (2003). Lipase catalysed regioselective esterification of protocatechuic aldehyde. Indian J. Chem. Section B. *42B*, 1119–1122.

Divakar, S. (2004). Porcine pancreas lipase catalysed ring-opening polymerization of ε - caprolactone. Journal of Macromolecular Science Part A – Pure and Applied Chemistry. A*41*, 537–546.

Dörmö, N., Bélafi-Bakó, K., Bartha, L., Ehrenstein, U. and Gubicza, L. (2004). Manufacture of an environmental-safe bio lubricant from fusel oil by enzymatic esterification in solvent-free system. Biochem. Eng. J. *21*, 229–234.

Engel, K.H., Bohnen, M. and Dobe, M. (1991). Lipase-catalysed reactions of chiral hydroxyacid esters: competition of esterification and transesterification. Enzyme Microb. Technol. *13*, 655–660.

Ferrer, M., Soliveri, J., Plou, F.J., López-Cortés, N., Reyes-Duarte, D., Christensen, M., Copa-Patiño, J.L. and Ballesteros, A. (2005). Synthesis of sugar esters in solvent mixtures by lipases from *Thermomyces lanuginosus* and *Candida antarctica* B and their antimicrobial properties. Enzyme Microb. Technol. *36*, 391–398.

From, M., Adlercreutz, P. and Mattiasson, B. (1997). Lipase catalyzed esterification of lactic acid. Biotechnol. Lett. *19*, 315–317.

Gayot, S., Santarelli, X. and Coulon, D. (2003). Modification of flavonoid using lipase in non-conventional media: effect of the water content. J. Biotechnol. *101*, 29–36.

Ghamgui, H., Karra-Chabouni, M. and Gargouri, Y. (2004). 1-Butyl oleate synthesis by immobilized lipase from *Rhizopus oryzae*: a comparative study between *n*-hexane and solvent-free system. Enzyme Microb. Technol. *35*, 355–363.

Gill, I. and Valivety, R. (1997). Polyunsaturated fatty acids: Part 1. Occurrence, biological activities and applications. Trends Biotechnol. *15*, 401–409.

Goto, M., Hatanaka, C. and Masahiro, G. (2005). Immobilization of surfactant–lipase complexes and their high heat resistance in organic media. Biochem. Eng. J. *24*, 91–94.

Hamsaveni, D.R., Prafulla, S.G. and Divakar, S. (2001). Optimization of isobutyl butyrate synthesis using central composite rotatable design. Process Biochem. *36*, 1103–1109.

Han, Y. and Chu, Y. (2005). The catalytic properties and mechanism of cyclohexane/DBSA/water microemulsion system for esterification. J. Mol. Catal. A Chem. *237*, 232–237.

Henderson, L.A., Svirikin, Y.Y., Gross, R.A., Kaplan, D.C. and Swift, G. (1996). Enzyme catalyzed polymerization of ε-caprolactone: effect of initiator on propagation. Polym. Mat. Sci. Eng. *75*, 6–7.

Hirakawa, H., Kamiya, N., Kawarabayashi, Y. and Nagamune, T. (2005). Log P effect of organic solvents on a thermophilic alcohol dehydrogenase. Biochim. Biophys. Acta. *1748*, 94–99.

Huang, S.Y., Chang, H.L. and Goto, M. (1998). Preparation of surfactant-coated lipase for the esterification of geraniol and acetic acid in organic solvents. Enzyme Microb. Technol. *22*, 552–557.

Ison, A.P., Dunill, P., Lilly, M., Macrae, A.R. and Smith, C. (1990). Enzymic interesterification of fats: immobilization and immunogold localization of lipase on ion-exchange resins. Biocatalysis *3*, 329–342.

Jeon, G.J., Park, O.J., Hur, B.K. and Yang, J.W. (2001). Enzymatic synthesis of amino acid- sugar alcohol conjugates in organic media. Biotechnol. Lett. *23*, 929–934.

Kanwar, L. and Goswami, P. (2002). Isolation of a *Pseudomonas* lipase produced in pure hydrocarbon substrate and its application in the synthesis of isoamyl acetate using membrane-immobilized lipase. Enzyme Microb. Technol. *31*, 727–735.

Khaled, N., Montet, D., Pina, M. and Graille, J. (1991). Fructose oleate synthesis in a fixed catalyst bed reactor. Biotechnol. Lett. *13*, 167–172.

Kim, J., Haam, S., Park, D.W., Ahn, I.S., Lee, T.G., Kim, H.S. and Kim, W.S. (2004). Biocatalytic esterification of β-methylglucoside for synthesis of biocompatible sugar-containing vinyl esters. Chem. Eng. J. *99*, 15–22.

Kiran, K.R. and Divakar, S. (2001). Lipase catalysed esterification of organic acids with lactic acid. J. Biotechnol. *87*, 109–121.

Kiran, K.R. and Divakar, S. (2003). Lipase catalysed polymerisation of lactic acid. World J. Microbiol. Biotechnol. *19*, 859–865.

Kiran, K.R., Karanth, N.G. and Divakar, S. (2002). Hydrogen ion concentration at the microaqueous phase in lipase catalysed esterification in non-aqueous organic media - steroyllactic acid. Indian J. Biochem. Biophys. *39*, 101–105.

Kiran, K.R., Suresh-Babu, C.V. and Divakar, S. (2001a). Thermostability of porcine pancreas lipase in non-aqueous media. Process Biochem. *36*, 885–892.

Kiran, K.R., Manohar, B. and Divakar, S. (2001b). A central composite rotatable design analysis of lipase catalysed synthesis of lauroyl lactic acid at bench-scale level. Enzyme Microb. Technol. *29*, 122–128.

Kittleson, J.R. and Pantaleone. (1994). Enzymic biphasic process for the synthesis of aromatic esters flavoring agents from corresponding carboxylic acid and alcohol by esterification mediated by a lipase from *Candida cyclindracea*. U.S. Patent. 5,437,991.

Klibanov, A.M. (1986). Enzymes that work in organic solvents. Chem. Technol. *16*, 354–359.

Knez, Z., Leitgeb, M., Zavrsnik, D. and Lavrie, B. (1990). Synthesis of oleic acid esters with immobilized lipase. Fat Sci. Technol. *4*, 169–172.

Kobayashi, S., Uyama, H., Namekawa, S. and Hayakawa, H. (1998). Enzymatic ring-opening polymerization and copolymerization of 8-octanolide by lipase catalyst. Macromolecules *31*, 5655–5659.

Krieger, N., Bhatnagar, T., Baratti, J.C., Baron, A.M., De-Lima, V.M. and Mitchell, D. (2004). Non-aqueous biocatalysis in heterogeneous solvent systems. Food Technol. Biotechnol. *42*, 279—286.

Kumar, R., Modak, J. and Madras, G. (2005). Effect of the chain length of the acid on the enzymatic synthesis of flavors in supercritical carbon dioxide. Biochem. Eng. J. *23*, 199–202.

Kumura, H., Mikawa, K. and Saito, Z. (1993). Influence of milk proteins on the thermostability of the lipase from *Pseudomonas fluorescence* 33. J. Dairy Sci. *76*, 2164–2167.

Kung, S. and Rhee, J. (1989). Effect of solvents on hydrolysis of olive oil by immobilized lipase in reverse phase system. Biotechnol. Lett. *11*, 37–42.

Laane, C., Boeren, S., Vos, K. and Veeger, C. (1987). Rules for optimization of biocatalysis in organic solvents. Biotechnol. Bioeng. *30*, 81–87.

Langrand, G., Rondot, N., Triantaphylides, C. and Baratti, J. (1990). Short-chain flavor esters synthesis by microbial lipases. Biotechnol. Lett. *12*, 581–586.

Leszczak, J.P. and Tran-Minh, C. (1998). Optimized enzymatic synthesis of methyl benzoate in organic medium. Operating conditions and impact of different factors on kinetics. Biotechnol. Bioeng. *60*, 556–561.

Lohith, K. and Divakar, S. (2005). Lipase catalysed synthesis of L-phenylalanine esters of D-glucose. J. Biotechnol. *117*, 49–56.

Ma, L., Persson, M. and Adlercreutz, P. (2002). Water activity dependence of lipase catalysis in organic media explains successful transesterification reactions. Enzyme Microb. Technol. *31*, 1024–1029.

Manohar, B. and Divakar, S. (2002). Application of central composite rotatable design to lipase catalyzed syntheses of m-cresyl acetate. World J. Microbiol. Biotechnol. *18*, 745–751.

Manohar, B. and Divakar, S. (2004a). Application of surface plots and statistical designs to selected lipase catalysed esterification reactions. Process Biochem. *39*, 847–851.

Manohar, B. and Divakar, S. (2004b). Porcine pancreas lipase acetylation of beta–cyclodextrin anchored 4-t-butylcyclohexanol. Indian J. Chem. Section B. *43B*, 2661–2665.

Marlot, C., Langrand, G., Triantaphylides, C. and Baratti, J. (1985). Ester synthesis in organic solvent catalyzed by lipase immobilized on hydrophilic supports. Biotechnol. Lett. *7*, 647–650.

Maruyama, T., Nagasawa, S.I. and Goto, M. (2002). Enzymatic synthesis of sugar esters in organic solvents. J. Biosci. Bioeng. *94*(4), 357–361.

Mestri, S. and Pai, J.S. (1994a). Effect of moisture on lipase catalyzed esterification of geraniol palmrosa oil in non-aqueous system. Biotechnol. Lett. *17*, 459–461.

Mestri, S. and Pai, J.S. (1994b). Synthesis of isoamyl butyrate by lipase of *Mucor miehei*. PAFAI J. *2*, 24–26.

Mishio, T., Takahashi, K., Yoshimoto, T., Kodera, Y., Saito, Y. and Inada, Y. (1987). Terpene alcohol ester synthesis by poylethylene glycol modified lipase in benzene. Biotechnol. Lett. *9*, 187–190.

Nagayama, K., Yamasaki, N. and Imai, M. (2002). Fatty acid esterification catalyzed by *Candida rugosa* lipase in lecithin microemulsion-based organogels. Biochem. Eng. J. *12*, 231–236.

Namekawa, S., Uyama, H. and Kobayashi, S. (1996). Lipase catalysed ring-opening polymerization and copolymerisation of β-propiolactone. Polym. J. *28*, 730–731.

Naoe, K., Ohsa, T., Kawagoe, M. and Imai, M. (2001). Esterification by *Rhizopus delemar* lipase in organic solvent using sugar ester reverse micelles. Biochem. Eng. J. *9*, 67–72.

Noel, M. and Combes, D. (2003). Effects of temperature and pressure on *Rhizomucor miehei* lipase stability. J. Biotechnol. *102*, 23–32.

Nogales, J.M.R., Roura, E. and Contreras, E. (2005). Biosynthesis of ethyl butyrate using immobilized lipase: a statistical approach. Process Biochem. *40*, 63–68.

Noureddini, H., Gao, X. and Philkana, R.S. (2005). Immobilized Pseudomonas cepacia lipase for biodiesel fuel production from soybean oil. Biores. Technol. *96*, 769–777.

Okumura, S., Iwai, M. and Tsujisaka, Y. (1979). Synthesis of various kinds of esters by four microbial lipases. Biochim. Biophys. Acta *575*, 156–165.

Osorio, N.M., Ferreira-Dias, S., Gusmao, J.H. and Da-Fonseca, M.M.R. (2001). Response surface modelling of the production of ω-3 polyunsaturated fatty acids-enriched fats by a commercial immobilized lipase J. Mol. Catal. B Enzym. *11*, 677–686.

Palomo, J.M., Segura, R.L., Mateo, C., Terreni, M., Guisan, J. M. and Fernandez-Lafuente, R. (2003). Synthesis of enantiomerically pure glycidol via a fully enantioselective lipase-catalyzed resolution. Enzyme Microb. Technol. *33*, 97–103.

Park, O.J., Park, H.G. and Yang, J.W. (1996). Enzymatic transesterification of monosaccharides and amino acid esters in organic solvents. Biotechnol. Lett. *18*, 473–478.

Park, O.J., Jeon, G.J. and Yang, J.W. (1999). Protease catalysed synthesis of disaccharide amino acid esters in organic media. Enzyme Microb. Technol. *25*, 455–462.

Perraud, R. and Laboret, F. (1989). Optimization of methyl propionate production catalyzed by *Mucor miehei* lipase. Appl. Microbiol. Biotechnol. *44*, 321–326.

Rahman, M.B.A., Md-Tajudin, S., Hussein, M.Z., Rahman, R.N.Z. R.A., Salleh, A.B. and Basri, M. (2005). Application of natural kaolin as support for the immobilization of lipase from *Candida rugosa* as biocatalsyt for effective esterification. Appl. Clay Sci. *29*, 111–116.

Rantwijk, F. and Sheldon, R.A. (2004). Enantioselective acylation of chiral amines catalysed by serine hydrolases. Tetrahedron *60*, 501–519.

Rao, P. and Divakar, S. (2001). Lipase catalysed esterification of α - terpineol with various organic acids application of the Plakett- Burman design. Process Biochem. *36*, 1125–1128.

Rao, P. and Divakar, S. (2002). Response surface methodological approach for the *Rhizomucor miehei* lipase- mediated esterification of α- terpineol with propionic acid and acetic anhydride. World J. Microbiol. Biotechnol. *18*, 341–345.

Rees, G.D. and Robinson, B.H. (1995). Esterification reactions catalyzed by *Chromobacterium viscosum* lipase in CTAB-based micro-emulsion systems. Biotechnol. Bioeng. *45*, 344–355.

Riva, S., Chopineau, J., Kieboom, A.P.G. and Klibanov, A.M. (1988). Protease catalysed regioselective esterification of sugars and related compounds in anhydrous dimethylformamide. J. Am. Chem. Soc. *110*, 584–589.

Romero, M.D., Calvo, L., Alba, C., Habulin, M., Primozic, M. and Knez, Z. (2005a). Enzymatic synthesis of isoamyl acetate with immobilized *Candida antarctica* lipase in supercritical carbon dioxide. J. Supercrit. fluids. *3*, 77–84.

Sakurai, T., Margolin, A.L., Russell, A.J. and Klibanov, A.M. (1988). Control of enzyme enantioselectivity by the reaction medium. J. Am. Chem. Soc. *110*, 7236–7237.

Schrag, J.D., Li. Y., Wu, S. and Cygler, M. (1996). Ser-His-Glu triad forms the catalytic site of the lipase from *Geotrichum candidum*. Nature *351*, 761–764.

Schreier, P. (1997). Biotechnology of aroma compounds. Berger RG (ed.). Adv. Biochem. Eng. Biotechnol. *17*, 52.

Segel, I.H. (1975). Enzyme Kinetics. John-Wiley and Sons, NY, USA.

Shaw, J.F., Wu, H.Z. and Shieh, C.J. (2003). Optimized enzymatic synthesis of propylene glycol monolaurate by direct esterification. Food Chem. *81*, 91–96.

Sheih, C.J., Akoh, C.C. and Yee, L.N. (1996). Optimized enzymatic synthesis of geranyl butyrate with lipase AY from *Candida rugosa*. Biotechnol. Bioeng. *51*, 371–374.

Shieh, C.J., Liao, H.F. and Lee, C.C. (2003). Optimization of lipase-catalyzed biodiesel by response surface methodology. Biores. Technol. *88*, 103–106.

Soo, E., Salleh, A.B., Basri, M., Rahman, R.N.Z.R.A. and Kamaruddin, K. (2003). Optimization of the enzyme-catalyzed synthesis of amino acid-based surfactants from palm oil fractions. J. Biosci. Bioeng. *95*, 361–367.

Stamatis, H., Xenkis, A., Menge, V. and Kolisis, N.F. (1993). Kinetic study of lipase catalyzed esterification in microemulsion. Biotechnol. Bioeng. *42*, 931–937.

Stamatis, H., Xenakis, A. and Kolisis, F.N. (1999). Bioorganic reactions in microemulsions: the case of lipases. Biotechnol. Adv. *17*, 293–318.

Suresh-Babu, C.V. and Divakar, S. (2001). Selection of alcohols through Plakett-Burman design in lipase catalyzed syntheses of anthranilic acid. J. Am. Oil Chem. Soc. *78*, 49–52.

Suresh-Babu, C.V., Karanth, N.G. and Divakar, S. (2002). Lipase catalysed esterification of cresols. Indian J. Chem. Section B. *41B*, 1068–1071.

Suzuki, Y., Shimizu, T., Takeda, H. and Kanda, K. (1991). Fermentative or enzymatic manufacture of sugar amino acid esters. Japan Patent. 03216194 A2.

Svirkin, Y.Y., Xu, J., Gross, R.A., Kaplan, D.L. and Swift, G. (1996). Enzymatic catalyzed stereoselective ring-opening polymerization of α-methyl-β-propiolactone. Macromolecules *29*, 4591–4597.

Torres, C. and Otero, C. (1999). Part 1: Enzymatic synthesis of lactate and glycolate esters of fatty alcohols. Enzyme Microb. Technol. *25*, 745–752.

Tramper, J., Vermie, M.H., Beetink, H.H. and Von-Stocker, U. (1992). In: Biocatalysis in non-conventional media. Elsevier, Amsterdam.

Turner, N.A., Duchateau, D.B. and Vulfson, E.N. (1995). Effect of hydration on thermostability of serine esterases. Biotechnol. Lett. *17*, 371–376.

Ulbrich, R., Golbik, R. and Schelleberger, A. (1991). Protein adsorption and leakage in carrier-enzyme systems. Biotechnol. Bioeng. *37*, 280–287.

Uyama, H., Kazuhiro, T., Norio, H. and Shiro, K. (1995). Lipase-catalyzed ring-opening polymerization of 12-dodecanolide. Macromolecules *28*, 7046–7050.

Uyama, H. and Kobayashi, S. (1996). Enzymatic ring-opening polymerization of macrolides to polyesters. Front. Biomed. Biotechnol. *3*, 5–15.

Valivety, R.H., Johannes, L.L., Rakels, L., Blanco, R.M., Johnston, R.M., Brown, L., Suckling, C.J. and Halling, P.J. (1990). Measurement of pH changes in an inaccessible aqueous phase during biocatalysis in organic media. Biotechnol. Lett. *12*, 475–480.

Vijayakumar, G.R., Lohith, K., Somashekar, B.R. and Divakar, S. (2004). Lipase catalysed synthesis of L-alanyl, L-Leucyl and L-phenylalanyl esters of D-glucose using unprotected amino acids. Biotechnol. Lett. *26*, 1323–1328.

Welsh, F.W., Williams, R.E. and Dawson, K.H. (1990). Lipase-mediated synthesis of low molecular weight flavor esters. J. Food Sci. *55*, 1679–1682.

Xu, Y., Wang, D., Qing Mu, X., Ao-Zhao, G. and Zhang, K.C. (2002). Biosynthesis of ethyl esters of short-chain fatty acids using whole-cell lipase from *Rhizopus Chinesis* CCTCC M201021 in non-aqueous phase. J. Mol. Catal. B Enzym. *18*, 29–37.

Yadav, G.D. and Devi, K.M. (2004). Immobilized lipase-catalysed esterification and transesterification reactions in non-aqueous media for the synthesis of tetrahydrofurfuryl butyrate: comparison and kinetic modeling. Chem. Eng. Sci. *59*, 373–383.

Yadav, G.D. and Lathi, P.S. (2003). Kinetics and mechanism of synthesis of butyl isobutyrate over immobilized lipases. Biochem. Eng. J. *16*, 245–252.

Ye, P., Xu, Z.K., Wang, Z.G., Wu, J., Deng, H.T. and Seta, P. (2005). Comparison of hydrolytic activities in aqueous and organic media for lipases immobilized on poly (acrylonitrile-*co*-maleic acid) ultrafiltration hollow fiber membrane J. Mol. Catal. B Enzym. *32*, 115–121.

Zaidi, A., Gainer, J.L., Carta, G., Mrani, A., Kadiri, T., Belarbi, Y. and Mir, A. (2002). Esterification of fatty acids using nylon-immobilized lipase in *n*-hexane: kinetic parameters and chain length effects. J. Biotechnol. *93*, 209–216.

Zaks, A. and Klibanov, A.M. (1986). Substrate specificity of enzymes in organic solvents vs. water is reversed. J. Am. Chem. Soc. *108*, 2767–2768

Zhang, T., Yang, L. and Zhu, Z. (2005). Determination of internal diffusion limitation and its macroscopic kinetics of the transesterification of CPB alcohol catalyzed by immobilized lipase in organic media. Enzyme Microb. Technol. *36*, 203–209.

Zhou, G.W., Li, G.Z., Xu, J. and Sheng, Q. (2001). Kinetic studies of lipase-catalyzed esterification in water-in-oil microemulsions and the catalytic behavior of immobilized lipase in MBGs. Colloids Surf. A: Physicochem. Eng. Aspects. *194*, 41–47.

CHAPTER 18

USE OF LIPASES IN ORGANIC SYNTHESIS

VICENTE GOTOR-FERNÁNDEZ AND VICENTE GOTOR*

Departamento de Química Orgánica e Inorgánica, Instituto Universitario de Biotecnología, Universidad de Oviedo, Oviedo, Spain

**vgs@fq.uniovi.es*

1. INTRODUCTION

The use of enzymes for organic synthesis has become an interesting area for organic and bio-organic chemists. Since many enzymes have been demonstrated to possess activity against non-natural substrates in organic media they have become widely used to carry out synthetic transformations. Hydrolases are the most frequently used enzymes due to their broad substrate spectrum and considerable stability. Additionally, many of them are commercially available and they work under mild reaction conditions and without the necessity for cofactors. Among the hydrolases, lipases (EC 3.1.1.3) are considered the most popular and useful enzymes for asymmetric synthesis (Wiktelius, 2005). Applications for lipases include kinetic resolution of racemic alcohols, acids, esters or amines (Ghanem and Aboul-Enein, 2004), as well as the desymmetrization of prochiral compounds (García-Urdiales *et al.*, 2005). They are also successfully employed in regioselective esterification or transesterification of polyfunctional compounds, for instance in the chemoenzymatic synthesis of nucleoside derivatives (Ferrero and Gotor, 2000). Recently, non-conventional processes, such as aldol reactions or Michael addition have been achieved using lipases (Bornscheuer and Kazlauskas, 2004).

2. LIPASES IN ORGANIC SOLVENTS

One of the more serious drawbacks for the use of enzymes in aqueous organic synthesis is the poor water solubility of organic compounds with more of four carbon atoms. Water is also a poor solvent for most applications in industrial chemistry, since many organic compounds are unstable in aqueous solution. Furthermore,

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 301–315

its removal is more tedious and expensive than that of organic solvents the boiling points of which are lower. The use of organic solvents presents several advantages including: (a) easier recovery of products with high yields; (b) the possibility of using non-polar substrates; (c) organic solvents avoid many side reactions, (d) in many cases lipases are thermodynamically more active than in water; (e) shifting the thermodynamic equilibrium to favour synthesis over hydrolysis.

Biocatalysis in non-aqueous media has been widely used for the resolution of alcohols, acids or lactones by enzymatic transesterification reactions using different lipases (Klibanov, 2001). Moreover, other processes such as the enzymatic acylation of amines or ammonia have shown themselves to be of great utility for the resolution of amines and the preparation of chiral amides.

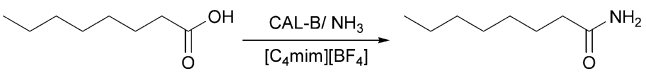
The main difference between the enzymatic acylation of alcohols and amines in organic solvents is the use of the corresponding acyl donor since activated esters which are useful in the enzymatic acylation of alcohols, such as halogen-ethyl or methyl esters, oxime ester, anhydrides and especially vinyl esters cannot be used with amines because they normally react in the absence of a biocatalyst. Thus non-activated esters must be used to carry out enzymatic aminolysis or ammonolysis reactions.

Although the enzymatic acylation of alcohols or amines is the process that has been most exhaustively studied, enzymatic alkoxycarbonylation for the activation and protection of hydroxyl or amino groups is a process of great utility. In this type of processes benzyl or allyl carbonates together with oxime or vinyl carbonates are the most efficient alkoxycarbonylating reagents (Takayama *et al*, 1999).

3. LIPASES IN IONIC LIQUIDS

The use of ionic liquids has recently emerged in organic synthesis and in some cases can be of great utility in biocatalysis (Jain *et al*, 2005). It constitutes an alternative for carrying out processes that present serious difficulties in organic solvents or water (Irimescu and Kato, 2004). These solvents can be used in three different ways in an enzymic system: (a) as a cosolvent in the aqueous phase, (b) as a pure solvent, and (c) as a two-phase system together with other solvents. Another possibility in biocatalysis is to carry out reactions under supercritical conditions. Although some progress has been achieved in this field, a major inconvenience is the need for sophisticated equipment. Nevertheless, this 'Green Chemistry' alternative could be of particular interest in the future.

One of the first examples of the use of ionic liquids was the synthesis of octanamide (Madeira Lau *et al*, 2000) by the reaction of octanoic acid and ammonia using *Candida antarctica* lipase B (CAL-B) as the biocatalyst (Scheme 1). Enzymatic aminolysis of carboxylic acids is more difficult to accomplish than that of esters because of the tendency of the reactants to form unreactive salts, however using carboxylic acids in the presence of ionic liquids the process takes place with high yields.



Scheme 1. Enzymatic preparation of octanamide in ionic liquids

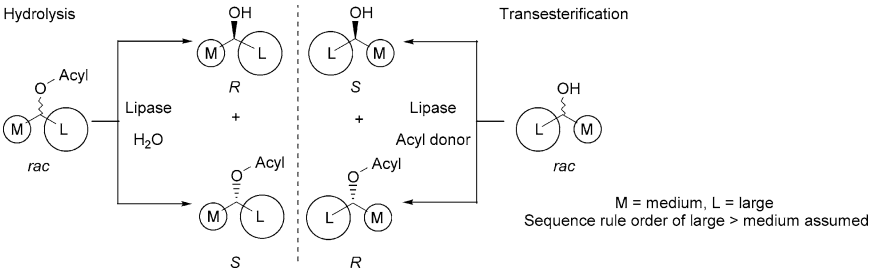
4. HYDROLYSIS OF ESTERS AND CARBONATES

Nowadays enzymatic transesterification is a process that is widely used for the preparation of chiral compounds. However, enzymatic hydrolysis, the natural reaction of lipases, is also very useful for the resolution of racemic mixtures or the desymmetrization of prochiral compounds. Hydrolysis and transesterification can be complementary processes for the resolution of secondary alcohols (Scheme 2). Both reactions fit Kazlauskas' rule (Kazlauskas *et al*, 1991) and the (*R*)-isomer reacts in both processes faster than its (*S*)-counterpart.

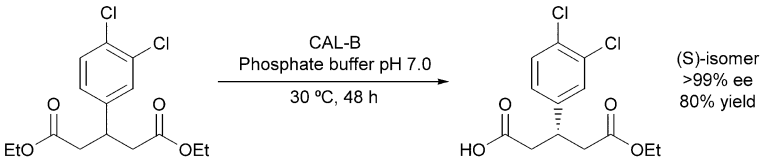
There are a number of enzymatic hydrolysis reactions that are of great utility for the synthesis of pharmaceuticals or their intermediates (Gotor, 2002). For example, β-adrenergic blocking agents such as propanolol, have been synthesized by chemoenzymatic methods, where the key step is an enzymatic hydrolysis. The main reason for preparing these amino alcohols in optically pure form is that the activity of these pharmaceuticals resides in the (*S*)-enantiomer. In addition, the preparation of single enantiomers of arylpropionic acids (non-steroidal anti-inflammatory agents) and optically pure 1,4-dihydropyridine derivatives (calcium antagonists) are important issues in the pharmaceutical industry.

The enzymatic desymmetrization of the prochiral compound diethyl 3-(3′, 4′- dichlorophenyl)-glutarate (Homann *et al*, 2001), an intermediate in the synthesis of neurokinin receptor antagonists, has been successfully developed and scaled up by enzymatic hydrolysis with CAL-B, obtaining the acid-ester enantiopure with 80% yield (Scheme 3).

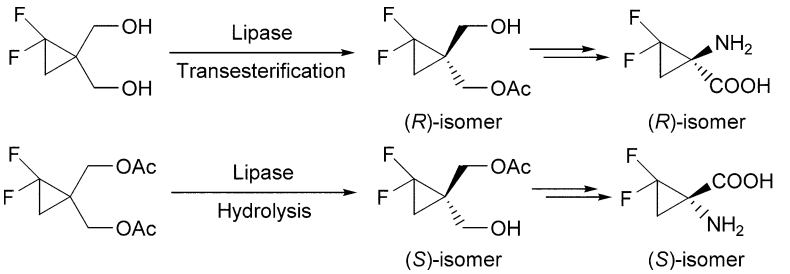
Hydrolysis and transesterification desymmetrization processes are in several cases complementary reactions of utility in obtaining both enantiomers of a compound with very high enantioselectivity. The synthesis of (*R*)- and (*S*)-1-amino-2,2-difluorocyclopropanecarboxylic acids via lipase-catalysed desymmetrization of



Scheme 2. Symmetry in enzymatic hydrolysis and transesterification



Scheme 3. Hydrolytic desymmetrization of 3-substituted glutarates



Scheme 4. Symmetry of hydrolysis and acylation in enzymatic desymmetrization

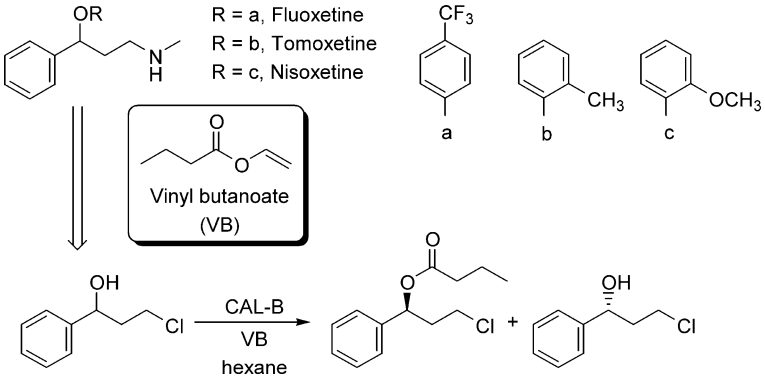
prochiral diols and diacetates has been reported (Kirihara *et al*., 2003). In both cases the products were isolated with very high yield and enantiomeric excesses (Scheme 4).

Although enzymatic ester hydrolysis has been widely applied in enzymatic resolution or desymmetrization processes, the enzymatic hydrolysis of carbonates has been scarcely reported. A practical example that we have carried out is the chemoenzymatic synthesis of (*S*)-zoplicone, a hypnotic that has a pharmaceutical profile of high efficacy and low toxicity. This was achieved by the hydrolysis of carbonates using lipases as biocatalysts (Fernández-Solares *et al*., 2002).

5. ESTERIFICATION AND TRANSESTERIFICATION PROCESSES

Enzymatic transesterification processes are more widely used than esterification reactions in resolution or desymmetrization processes, and the acylation of alcohols using lipases is currently the most frequently used process in biocatalysis (Faber, 2004). Whilst several agents can be employed, the most efficient are vinyl esters, especially vinyl acetate which in many cases is used as both solvent and acyl donor. There are many examples of the resolution of primary and secondary alcohols. Although the reaction with tertiary alcohols is more difficult some examples have been reported (Krishna *et al*., 2002).

Here we choose a few representative examples of acylation of alcohols where enzymatic transesterification is the key step in the synthesis of chiral pharmaceuticals. Fluoxetine, tomoxetine and nisoxetine, three antidepressants, were synthesised from racemic 3-chloro-1-phenylpropan-1-ol (Liu *et al*., 2000) by enzymatic

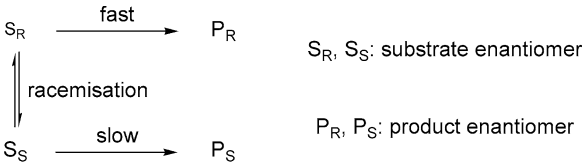


Scheme 5. Enzymatic transesterification of pharmaceutical intermediates

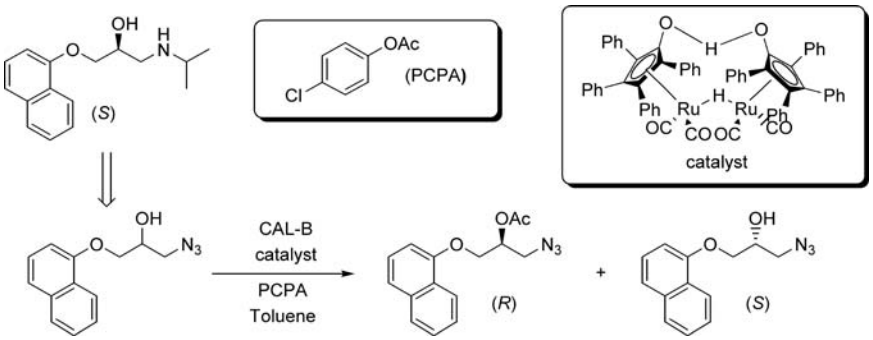
transesterification in hexane using vinyl butanoate (VB) as the acyl donor and CAL-B as the best biocatalyst (Scheme 5). It is also possible to obtain these pharmaceuticals by resolution of 3-hydroxy-3-phenylpropanenitrile as the starting material using *Pseudomonas cepacia* lipase (PSL) and vinyl acetate (VA) as the acyl donor (Kamal *et al*., 2002).

Kinetic resolution (KR) is a useful method to obtain enantiomerically pure compounds but suffers the drawback that the maximum yield is 50% of the starting material. This limitation can be overcome via dynamic kinetic resolution (DKR), in which the slower reacting enantiomer racemises during the process (Scheme 6). This procedure can theoretically lead to a single product enantiomer with 100% yield.

DKR has appeared in asymmetric catalysis in the last decade as a common methodology involving a lipase as the biocatalyst and a metal-organic complex as the chemical catalyst (Pàmies and Bäckvall, 2003). This strategy has been widely used for the resolution of alcohols by enzymatic hydrolysis or transesterification processes. For instance, the synthesis of (*S*)-propanolol by DKR of the corresponding azidoalcohol using a combination of a ruthenium complex and CAL-B in toluene at 80 °C and *p*-chlorophenyl acetate (PCPA) as the acyl donor (Scheme 7) has been described. After 1 day the (*R*)-acetate was produced with > 99% ee and 94% conversion (86% isolated yield). The enzyme was recycled and used again for another DKR without any loss of activity (Pàmies and Bäckvall, 2001).



Scheme 6. Schematic representation of a DKR process

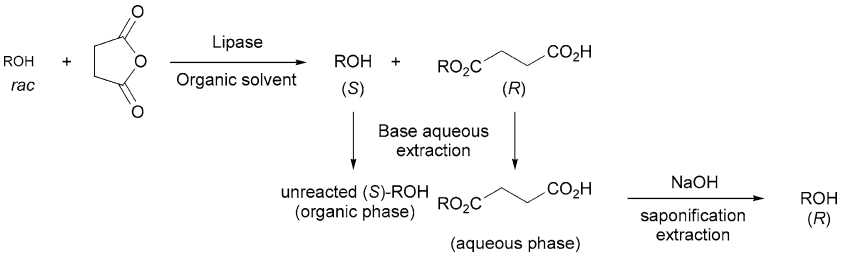


Scheme 7. Chemoenzymatic synthesis of (*S*)-propanolol by DKR

In some cases the use of vinyl or alkyl esters as acyl donors has the drawback of the need to separate the ester (product) from the alcohol (substrate). A practical strategy to avoid this problem is the use of cyclic anhydrides (Bouzemi *et al*., 2004). In this case an acid is obtained as the product which can be readily separated from the unreacted alcohol by a simple aqueous base-organic solvent liquid-liquid extraction (Scheme 8).

This strategy has been applied to the resolution of *N*-substituted *trans*-4-(4′-fluorophenyl)-3-hydroxymethylpiperidines as key intermediates in the synthesis of (−)-paroxetine which is a potent and selective inhibitor of 5-hydroxytryptamine reuptake and is used in the treatment of a variety of human diseases such as depression (De Gonzalo *et al*., 2003). The best results were obtained with a combination of CAL-B, glutaric anhydride and toluene as the solvent. In addition, the reuse of the immobilized lipase afforded the same enantioselectivity in the second and third cycles with just a moderate loss of enzyme activity in the fourth and fifth cycles.

Lipases are also valuable tools for the resolution of biaryl derivatives with axial chirality (Sanfilippo *et al*., 2003). Thus, *Pseudomonas cepacia* lipase (PSL) has been used for the transesterification of 2, 2′-dihydroxy-6, 6′-dimethoxy-1, 1′-biphenyl in a reaction using vinyl acetate as the acyl donor and *tert*-butyl methyl ether as



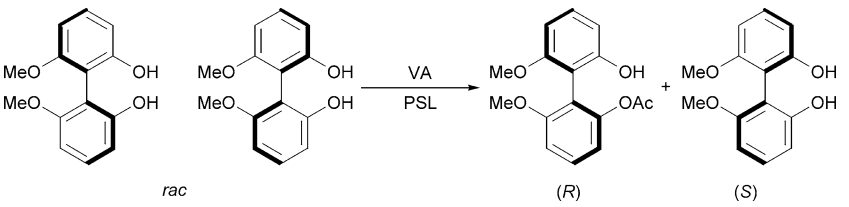
Scheme 8. Enzymatic resolution of alcohols using anhydrides as acyl reagents

the solvent. Product of configuration (*R*) was obtained with >98% ee while the (*S*)-substrate was recovered with >96% ee (Scheme 9). These compounds are of great interest as ligands in asymmetric organic synthesis.

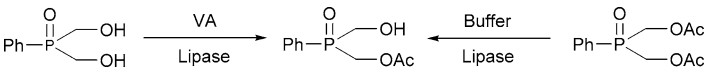
As commented above, in enzymatic desymmetrization processes it is possible to achieve an enantiopure compound with a maximum yield of 100%. For this reason these reactions constitute a very interesting alternative to KRs. One of the families of substrates to which more attention has been being paid is that containing the propane-1,3-diol moiety because this group is present or can easily lead to many molecules that play important roles in medicinal chemistry and/or asymmetric synthesis (Neri and Williams, 2003).

The first enzymatic desymmetrization of prochiral phosphine oxides has recently been published (Kielbasinski *et al*, 2003). In this case the chiral centre is the phosphorus atom. The prochiral compound bis(hydroxymethyl)phenyl-phosphine oxide was desymmetrized using either lipase-catalysed acetylation with vinyl acetate as acyl donor in an organic solvent, or enzymatic hydrolysis of the corresponding diacetate in phosphate buffer and solvent (Scheme 10). The monoacetate was obtained in 79% ee and 76% yield.

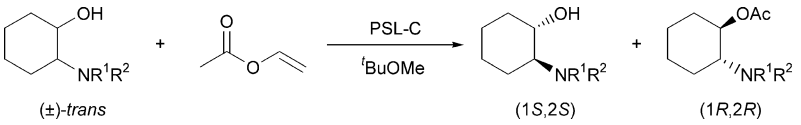
Chiral 1,2 and 1,3-amino alcohols have proven to be a functionally active class of compounds of wide application in medical chemistry. The *N*-acylation to obtain the corresponding amide gives very poor yields because migration of the acyl group normally takes place. For this reason, to achieve a good bioresolution with amino alcohols it is necessary to protect the amino or hydroxyl group. For instance, the four isomers of 1,2-aminoindanols can be resolved using carbamate derivatives. Of special relevance is the enzymatic transesterification of *cis*-(1*S*,2*R*)-1-aminoindan -2-ol (Luna *et al*, 1999), a key component of indinavir which is a potent inhibitor of the protease of the human immunodeficiency virus (HIV). Optically active *trans*-2-(*N*, *N*-dialkylamino)-cyclohexanols have been easily prepared in a two-step sequence: ring opening of cyclohexene oxide and subsequent resolution of the resulting racemic amino alcohol by transesterification catalysed by PSL-C



Scheme 9. Enzymatic resolution of a compound with axial chirality



Scheme 10. Enzymatic desymmetrization of prochiral phosphine oxide



Scheme 11. Enzymatic resolution of 1,2-amino alcohols

(González-Sabín *et al*., 2004). In most cases conversions reached values closed to 50%, isolating substrate and product in enantiopure form (Scheme 11). The utility of these β-amino alcohols as chiral ligands has also been investigated.

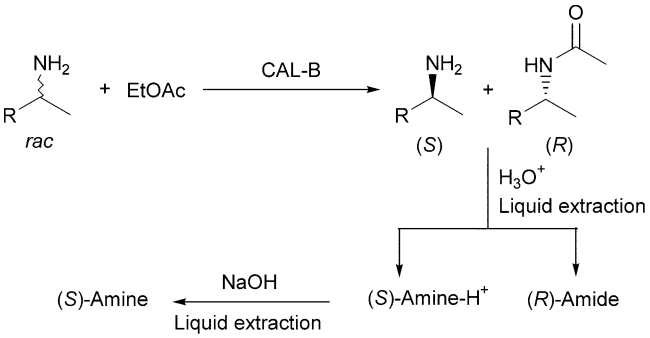
6. ENZYMATIC AMINOLYSIS AND AMMONOLYSIS REACTIONS

Traditional synthetic methods to obtain optically pure amines have used chiral catalysts for the reduction of amine precursors. However the preparation of enantio-enriched amines via lipase-catalysed enantioselective acylation can be accomplished using mild conditions, non-toxic reagents and easy experimental procedures, and it is also possible to recycle the biocatalyst. As a result, the use of enzymatic methods for the preparation of chiral nitrogenated compounds has rapidly gained prominence in Green Chemistry and, of course, in large-scale industrial applications. In addition, lipases are the most efficient hydrolases for catalysing the acylation of amines and ammonia because they have very low amidase activity.

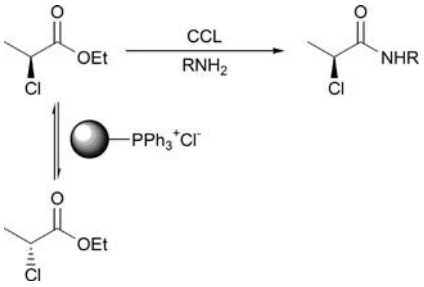
In the last few years several reviews have been published showing the utility of lipases in ammonolysis and aminolysis reactions for the synthesis of nitrogen-containing organic compounds (Gotor, 1999; van Rantwijk and Sheldon, 2004; Alfonso and Gotor, 2004). In this section we describe a few representative examples of enzymatic aminolytic and ammonolytic processes using lipases in organic solvents. The best biocatalyst for these enzymatic reactions is generally CAL-B, although in some cases CAL-A or PSL-C can also be used depending of the structure of the amine.

Scheme 12 shows the general strategy to obtain enantiopure amines (González-Sabín *et al.*, 2002). Normally ethyl acetate is used for the acylation of amines, in many cases as both acyl donor and solvent. Other acylating agents such alkyl methoxy acetates are also of utility, however vinyl esters, the best reagents for the resolution of alcohols, are not adequate for the resolution of primary amines due to their high reactivity. Although there are many examples of KR of primary amines (Alfonso and Gotor, 2004), few examples of the preparation of enantiomerically pure secondary amines by enzymatic acylation have been reported.

In some cases sequential biocatalytic resolutions by 'one-pot' double enzymatic reactions are of great utility because with moderate enantioselectivity in both processes it is possible to achieve a high ee of substrate and product. An example of this is the resolution of *trans*-cyclohexane-1,2-diamine and *trans*-cyclopentane-1,2-diamine (Alfonso *et al*., 1996; Luna *et al*., 2002).



Scheme 12. KR of primary amines by enzymatic acylation

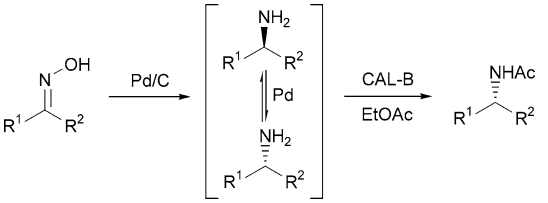


Scheme 13. DKR of an ester by aminolysis

Examples of DKR via enzymatic aminolysis reactions are scarce in the literature. Nevertheless, the resolution of racemic ethyl 2-chloropropionate amines using *Candida cylindracea* lipase (CCL) has been carried out in this way catalysed by encapsulated CCL in the presence of triphenylphosphonium chloride immobilized on Merrifield resin (Scheme 13). This process yielded the (*S*)-enantiomer with high yield and ee (Badjić *et al*, 2001).

An elegant example of the DKR of amines has been described using ketoximes as the starting material (Choi *et al*, 2001). The coupling of Pd-catalysed reduction of ketoximes and the subsequent Pd and CAL-B catalysed DKR of the resulting racemic amine afforded acetamides with very high chemical and optical yields. In this process an additive such as *N*-ethyldiisopropylamine was required to suppress the reductive deamination of the amine (Scheme 14). This procedure improves the results of the first DKR reported (Reetz and Schimossek, 1996). By this means the concentration of amine is low and the formation of by-products via reductive amination is less favoured.

The resolution of secondary amines is a process that presents more difficulty than the resolution of primary amines. Cyclic secondary amines are structurally easier to resolve than acyclic compounds and examples of the resolution of pyrrolidine and

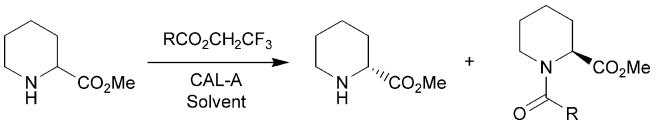


Scheme 14. DKR of racemic amines from ketoximes

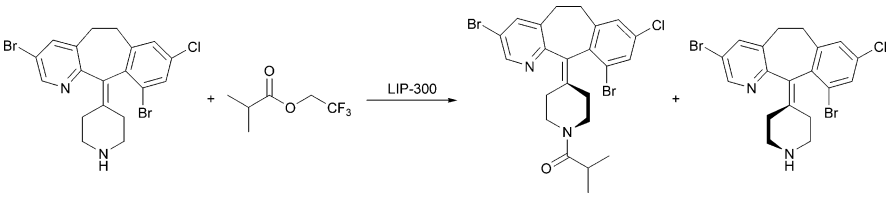
piperidine derivatives have been described. The enzymatic acylation of pipecolic acid derivatives has been catalysed by CAL-A with good results (Liljeblad *et al*, 2002). This enzyme seems to have a larger pocket in the active site than CAL-B and accepts bulkier substrates. This reaction has recently been applied to a pyrrolidine ring and the enzyme was very efficient in catalysing the acylation of the secondary amino group with very high enantioselectivity (Scheme 15).

The resolution of secondary amines via enzyme-catalysed acylation is not frequently used. In the case of piperidine shown in Scheme 16, the molecule exists as a pair of enantiomers due to atropisomerism about the exocyclic double bond. The use of the lipase Toyobo LIP-300 and trifluoroethyl isobutyrate as the acylating agent resulted in isobutyrylation of the (+)-enantiomer which was used as the starting material for the synthesis of a product of physiological interest (Morgan *et al.*, 2000).

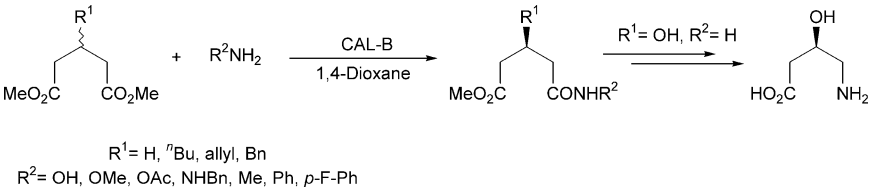
The first example of desymmetrization by enzymatic aminolysis and ammonolysis reactions was described several years ago (Puertas *et al*, 1996). The aminolysis of dimethyl 3-hydroxyglutarate with amines and ammonia in the presence of CAL-B led exclusively to the corresponding monoamide of configuration (*S*) (Scheme 17). The ammonia-derived enantiopure monoamide has been



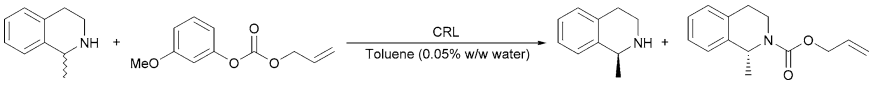
Scheme 15. Example of the resolution of secondary amines



Scheme 16. KR of atropisomers with lipases



Scheme 17. Synthesis of (*R*)-GABOB by ammonolytic desymmetrization



Scheme 18. Resolution of a secondary amine by enzymatic alkoxycarbonylation

used to prepare the biologically interesting β-amino acid (*R*)-3-hydroxy-4-aminobutanoic acid [(*R*)-GABOB]. This desymmetrization process has also been applied to other 3-substituted glutarates (López-García *et al*., 2003a), and some of these enantiopure monoamides have been used to prepare β-amino acids (López-García *et al*., 2003b).

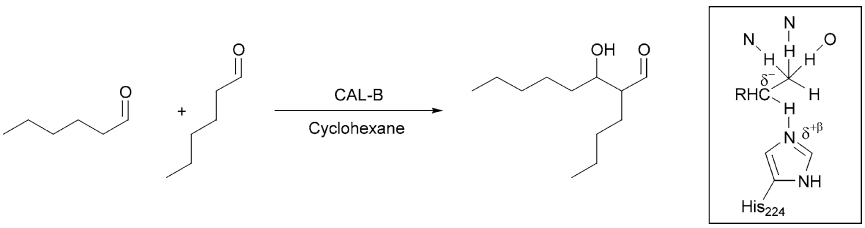
Recently enzymatic alkoxycarbonylation has been applied to resolve the secondary amine 1-methyl tetrahydroisoquinoline (Scheme 18) using substituted phenyl allylcarbonates as acylating reagents and *Candida rugosa* lipase (CRL) as biocatalyst (Breen, 2004). The best solvent was found to be toluene and the (*S*)-amine was recovered with 46% yield and 99% ee whilst the (*R*)-carbonate was obtained at 47% yield and 98% ee. To achieve these results the amount of water used in the process is critical.

7. NON-CONVENTIONAL REACTIONS OF LIPASES

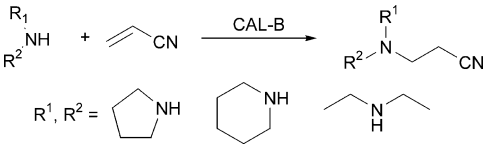
In many cases, enzymes are able to catalyse more than one reaction. The challenge is to use mechanistic reasoning to discover these new processes. Among the lipases CAL-B is the biocatalyst that has shown the greatest promiscuity (Kazlauskas, 2005). This author defines 'catalytic promiscuity' as the ability of a single active site to catalyse more than one chemical reaction. It has been reported that this lipase can catalyse aldol condensations and Michael additions.

Aldol condensation of hexanal in cyclohexane is catalysed by CAL-B (Branneby *et al.*, 2003). The reaction is not enantioselective and the authors hypothesized that the formation of a carbon-carbon bond did not require the active site serine. Indeed replacement with alanine increased the aldol addition approximately two-fold. The calculated transition-state structure for enolate formation is shown on the right in Scheme 19.

Two recent articles have noted the potential of CAL-B to catalyse Michael-type additions. This lipase catalyses the addition of thiols or secondary amines to



Scheme 19. Aldol condensation of hexanal catalysed by lipases



Scheme 20. CAL-B catalyses Michael additions

α, β-unsaturated aldehydes (Carlqvist *et al*., 2004). Quantum-modelling suggests that the oxyanion hole in CAL-B activates the aldehyde for addition, the histidine residue acts as a base, while serine is not involved in this process (Branneby *et al*., 2004).

The reaction of acrylonitrile with secondary amines in the presence of CAL-B led to the production of the corresponding Michael adduct faster than in the absence of biocatalyst (Scheme 20) and a tentative mechanism for this new process has been proposed (Torre *et al*., 2004).

A new strategy for the enzymatic synthesis of pyrimidine derivatives containing a sugar branch has been developed combining enzymatic Michael addition and acylation processes. The first step in the reaction between pyrimidines and vinyl 3-propionyloxy propionate was catalysed by Amano lipase M from *Mucor javanicus* in DMSO, while the regioselective acylation of D-glucose and D-mannose with the Michael adducts was catalysed by alkaline protease from *Bacillus subtilis* in pyridine (Xu *et al*., 2005).

8. CONCLUDING REMARKS

The use of lipases has become a conventional process in organic synthesis, not only for the preparation of optically pure compounds but also for regioselective and chemoselective processes. Their utility in carrying out selective transformations under mild reaction conditions make them attractive catalysts for performing certain transformations that are difficult to achieve by chemical procedures. Nowadays many companies use lipases for the preparation of chemicals instead of using chemical catalysis because the use of these biocatalysts has enormous advantages including the economy of the process, the environmental friendliness of the

catalysts and their recyclability. In addition, genetic engineering techniques can be expected to play a major role in future research providing new biocatalytic pathways ultimately leading to the generation of a great variety of new products.

REFERENCES

Alfonso, I., Astorga, C., Rebolledo, F. and Gotor, V. (1996). Sequential biocatalytic resolution of (±)-*trans*-cyclohexane-1,2-diamine. Chemoenzymic synthesis of an optically active polyamine. Chem. Commun. 2471–2472.

Alfonso, I. and Gotor, V. (2004). Biocatalytic and biomimetic aminolysis reactions: useful tools for selective transformations on polifunctional substrates. Chem. Soc. Rev. *33*, 201–209.

Badjić, J.D., Kadnikova E.N. and Kostić, N.M. (2001). Enantioselective aminolysis of an α-chloroester catalyzed by *Candida cylindracea* lipase encapsuled in sol-gel silica glass. Org. Lett. *3*, 2025–2027.

Breen, G.F. (2004). Enzymatic resolution of a secondary amine using novel acylating reagents. Tetrahedron: Asymmetry *15*, 1427–1430.

Bornscheuer, U.T. and Kazlauskas R.J. (2004). Catalytic promiscuity in biocatalysis: Using old enzymes to form new bonds and follow new pathways. Angew. Chem. Int. Ed. *43*, 6032–6040.

Bouzemi, N., Debbeche, H., Aribi-Zouiou, L. and Fiaud, J-C. (2004). On the use of succinic anhydride as acylating agent for practical resolution of aryl-alkyl alcohols through lipase-catalysed acylation. Tetrahedron Lett. *45*, 627–630.

Branneby, C., Carlqvist, P., Magnusson, A., Hult, K., Brinck, T. and Berglund, P. (2003). Carbon-Carbon bonds by hydrolytic enzymes. J. Am. Chem. Soc. *125*, 874–875.

Branneby, C., Carlqvist, P., Hult, K., Brinck, T. and Berglund, P. (2004). Aldol additions with mutant lipase: analysis by experiments and theoretical calculations. J. Mol. Catal. B: Enz. *31*, 123–128.

Carlqvist, P., Svedendahl M., Branneby, C., Hult, K., Brinck, T. and Berglund, P. (2004). Exploring the active-site of a rationally redesigned lipase for catalysis of Michael-type additions. ChemBioChem *6*, 331–336.

Choi, Y.K., Kim, M.J., Ahn, Y. and Kim M-J. (2001). Lipase/palladium-catalyzed asymmetric transformations of ketoximes to optically active amines. Org. Lett. *3*, 4099–4101.

De Gonzalo, G., Brieva, R., Sánchez, V., Bayod, M. and Gotor, V. (2003). Anhydrides as acylating agents in the enzymatic resolution of an intermediate of (−)-paroxetine. J. Org. Chem. *68*, 3333–3336.

Faber, K. Biotransformations in Organic Chemistry. Springer-Verlag: Heidelberg, 2004

Fernández-Solares, L., Díaz, M., Brieva, R., Sánchez, V.M., Bayod, M. and Gotor, V. (2002). Enzymatic resolution of new carbonate intermediates for the synthesis of (*S*)-(±)-zopiclone. Tetrahedron: Asymmetry *13*, 2577–2582.

Ferrero, M. and Gotor, V. (2000). Biocatalytic selective modifications of conventional nucleosides, carbocyclic nucleosides and C-nucleosides. Chem. Rev. *100*, 4319–4347.

García-Urdiales, E., Alfonso, I. and Gotor, V. (2005). Enantioselective desymmetrization in organic synthesis. Chem. Rev. *105*, 313–354.

Ghanem A. and Aboul-Enein H.Y. (2004). Lipase mediated chiral resolution of racemates in organic solvents. Tetrahedron: Asymmetry *15*, 3331–3351.

González-Sabín, J., Gotor, V. and Rebolledo, F. (2002). CAL-B-catalyzed resolution of some pharmacologically interesting β-substituted isopropylamines. Tetrahedron: Asymmetry *13*, 1315–1320.

González-Sabin, J., Gotor, V. and F. Rebolledo, F. (2004). Chemoenzymatic preparation of optically active β-aminocyclohexanols and their application in the enantioselecetive addition of diethylzinc to benzaldehyde. Tetrahedron: Asymmetry *15*, 1335–1341.

Gotor, V. (1999). Non-Conventional Hydrolase Chemistry: Amide and Carbamate Bond Formation Catalyzed by Lipases. Bioorg. Med. Chem. *7*, 2189–2197.

Gotor, V. (2002). Biocatalysis applied to the preparation of pharmaceuticals. Org. Proc. Res. Dev. *6*, 420–426.

Homann, M.J., Vail, R., Morgan, B., Sabesan, V., Levy, C., Dodds, D.R. and Zaks, A. (2001). Enzymatic Hydrolysis of a Prochiral 3-Substituted Glutarate Ester, an Intermediate in the Synthesis of an NK_1/NK_2 Dual Antagonist Adv. Synth. Catal. *343*, 744–749.

Irimescu, R. and Kato, K. (2004). Lipase-catalyzed enantioselective reaction of amines with carboxylic acids under reduced pressure in non-solvent system and in ionic liquids. Tetrahedron Lett. 45, 523–525.

Jain, N., Kumar, A., Chauhan, S. and Chauhan, S.M.S. (2005). Chemical and biochemical transformations in ionic liquids. Tetrahedron *61*, 1015–1060.

Kamal, A., Khanna, G.B.R. and Ramu, R. (2002). Chemoenzymatic synthesis of both enantiomers of fluoxetine, tomoxetine and nisoxetine: lipase catalysed resolution of 3-aryl-3-hydroxypropanenitriles. Tetrahedron: Asymmetry *13*, 2039–2051.

Kazlauskas, R.J., Weissfloch, A.N.E Rappaport A.T. and Cuccia, L.A. (1991). A rule to predict which enantiomer of a secondary alcohol reacts faster in reactions catalyzed by cholesterol esterase, lipase from *Pseudomonas cepacia*, and lipase from *Candida rugosa*. J. Org. Chem. *56*, 2656–2665.

Kazlauskas, R.J. (2005). Enhancing catalytic promiscuity for biocatalysis. Curr. Opin. Chem. Biol. *9*, 195–201.

Kielbasinski, P., Zurawinski, R., Albrycht, M. and Mikolajczyk, M. (2003). The first enzymatic desymmetrizations of prochiral phosphine oxides. Tetrahedron: Asymmetry *14*, 3379–3384.

Kirihara, M., Kawasaki, M., Takuwa, T., Kakuda, H., Wakikawa, T., Takeuchi, Y. and Kirk, K.L. (2003). Efficient synthesis of (*R*)- and (*S*)-1-amino-2,2-difluorocyclopropanecarboxylic acid via lipase-catalyzed desymmetrization of prochiral precursors. Tetrahedron: Asymmetry *14*, 1753–1761.

Klibanov, A.M. (2001) Improving enzymes by using them in organic solvents. Nature *409*, 241–246.

Krishna, S.H., Persson, M., Bornscheuer, U.T. (2002). Enantioselective transesterification of a tertiary alcohol by lipase A from *Candida antarctica*. Tetrahedron: Asymmetry *13*, 2693–2696.

Liljeblad, A., Lindborg, J., Kanerva, A., Katajisto, J. and Kanerva, L.T. (2002). Enantioselective lipase-catalyzed reactions of methyl pipecolinate: transesterification and *N*-acylation. Tetrahedron Lett. *43*, 2471–2474.

Liu, H.-L., Helge, B. and Anthonsen, T. Chemoenzymatic synthesis of the non-tricyclic antidepressants Fluoxetine, Tomoxetine and Nisoxetine. (2000). J. Chem. Soc. Perkin Trans. *1*, 1767–1769.

López-García, M., Alfonso, I., Gotor, V. (2003a). Desymmetrization of dimethyl 3-substituted glutarates through enzymatic ammonolysis and aminolysis reactions. Tetrahedron: Asymmetry *14*, 603–609.

López-García, M., Alfonso, I. and Gotor, V. (2003b). Synthesis of (*R*)-3,4-diaminobutanoic acid by desymmetrization of dimethyl 3-(benzylamino)glutarate through enzymatic ammonolysis. J. Org. Chem. *68*, 648–651.

Luna, A., Maestro, A., Astorga, C. and Gotor, V. (1999). Enzymatic resolution of (±)-*cis* and (±)-*trans*-1-aminoindan-2-ol and (±)-*cis* and (±)-*trans*-2-aminoindan-1-ol Tetrahedron: Asymmetry *10*, 1969–1977.

Luna, A., I. Alfonso, I. and Gotor, V. (2002). Biocatalytic approaches toward the synthesis of both enantiomers of *trans*-cyclopentane-1,2-diamine. Org. Lett. *4*, 3627–3629.

Madeira Lau, R., Rantwijk, F.V., Seddon, K.R. and Sheldon R.A. (2000). Lipase-catalyzed reaction in ionic liquids. Org. Lett. *2*, 4189–4191.

Morgan, B., Zaks, A., Dodds, D.R., Liu, J., Jain, R.,Megati, S., Njoroge F.J. and Girijavallabhan, V.M. (2000). Enzymatic kinetic resolution of piperidine atropoisomers: synthesis of a key intermediate of the farnesyl protein transferase inhibitor. J. Org. Chem. *65*, 5451–5459.

Neri, C. and Williams, J.M.J. (2003). New Routes to Chiral Evans Auxiliaries by Enzymatic Desymmetrisation and Resolution Strategies. Adv. Synth. Catal. *345*, 835–848.

Pàmies, O. and Bäckvall, J.-E. (2001). Dynamic kinetic resolution of β-azido alcohols. An efficient route to chiral aziridines and β-aminoalcohols. J. Org. Chem. *66*, 4022–4025.

Pàmies, O. and Bäckvall, J-E. (2003). Combination of enzymes and metal catalysis. A powerful approach in asymmetric catalysis. Chem. Rev. *103*, 3247–3262.

Puertas, S., Rebolledo, F., Gotor, V. (1996). Enantioselective enzymic aminolysis and ammonolysis of dimethyl 3-hydroxyglutarate. Synthesis of (*R*)-4-amino-3-hydroxybutanoic acid. J. Org. Chem. *61*, 6024–6027.

Reetz, M.T. and Schimossek, K. (1996). Lipase-catalyzed dynamic kinetic resolution of chiral amines: Use of palladium as the racemization catalyst. Chimia 50, 668–669.

Sanfilippo, C., Nicolosi, G., Fabbri, D. and Dettori, M.A. (2003). Access to optically active 2, 2′-dihydroxy-6, 6′-dimethoxy-1, 1′-biphenyl by a simple biocatalytic procedure. Tetrahedron: Asymmetry *14*, 3267–3270.

Takayama, S., Lee, S.T., Hung, S-H., Wong, C-H. (1999). Designing enzymic resolution of amines. Chem. Commun. 127–128.

Torre, O., Alfonso, I. and Gotor, V. (2004). Lipase catalysed Michael addition of secondary amines to acrylonitrile. Chem. Commun. 1724–1725.

van Rantwijk, F. and Sheldon, R.A. (2004). Enantioselective acylation of chiral amines catalysed by serine hydrolases. Tetrahedron *60*, 501–519.

Xu, J-M., Yao, S-P., Wu, W-B., Lv, D-S and Lin X-F. (2005). Two-step sequential synthesis of pyrimidine derivatives containing a sugar branch via combining of enzymatic Michael addition/acylation. J. Mol. Catal. B: Enz. *35*, 122–127.

Wiktelius, D. (2005). Lipases - Enzymes for biocatalytic asymmetric synthesis. Synlett 2111–2114.

CHAPTER 19

USE OF LIPASES FOR THE PRODUCTION OF BIODIESEL

ANDREA SALIS*, MAURA MONDUZZI AND VINCENZO SOLINAS
Dipartimento di Scienze Chimiche, Università di Cagliari - CSGI, Monserrato (CA), Italia
**asalis@unica.it*

1. INTRODUCTION

1.1. Why Biodiesel?

Biodiesel is composed of a mixture of fatty acid alkyl esters. It is a natural substitute for petroleum-derived diesel fuel and has similar or better specifications (density, viscosity, cetane number, flash point, *etc*.). Biodiesel is industrially obtained by transesterification of vegetable oils or animal fats with short chain alcohols. When methanol is used, the resulting biodiesel is a mixture of fatty acid methyl esters (FAME). Other alcohols such as ethanol, *iso*-propanol or longer linear or branched chains can also be used. Whilst methanol is the cheapest alcohol, the other alcohols yield products with better performances (Knothe, 2005a). Since biodiesel comes from renewable sources it is CO_2-neutral, biodegradable and conserves fossil fuels. Compared to traditional diesel fuel its combustion leads to a substantial reduction in polluting emissions. Finally, biodiesel is less dangerous to handle than diesel fuel because of its higher flash point (120 °C compared to 61 °C).

Interest in alternative energy sources is justified by high petroleum prices and increasing environmental concerns. Countries committed to the Kyoto protocol must decrease CO_2 and the other greenhouse gas emissions by 8% by 2012 – with respect to those measured in 1990. In this context the European Union has decided to increase the use of biofuels from 1.7 % in 2003 to 5.75 % of total diesel fuel consumption by 2010. This means that production must triplicate in the next few years, and justifies the major interest being given to this biofuel.

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 317–339

1.2. Disadvantages of Presently used Processes for Biodiesel Production

Fig. 1a shows the industrially used reaction for biodiesel production by homogeneous alkaline catalysis. The basic catalyst, *i.e.* sodium hydroxide, reacts with methanol to yield sodium methoxide (Fig. 1b) that then reacts with the triglyceride to produce the FAME. In order to increase the reaction rate, temperatures greater than 60–70 °C are needed, demanding high energy consumption. The reaction products are a mixture of esters and several by-products. The main by-product is glycerol that, after purification, can be used for cosmetic and pharmaceutical purposes. Other by-products are di- and mono-glycerides arising from the partial alcoholysis of triglycerides. These by-products must be separated and pumped back to the reactor. Free fatty acids, water and unreacted alkaline catalyst are also present. Thus complicated purification processes are needed in order to obtain a pure biodiesel achieving the standard requirements.

When the acidity of a feedstock is high the reaction between the free fatty acids and the basic catalyst produces soap (Fig. 1c). Since soap is a surfactant it forms emulsions and makes the separation between FAME and glycerol difficult. Thus, in

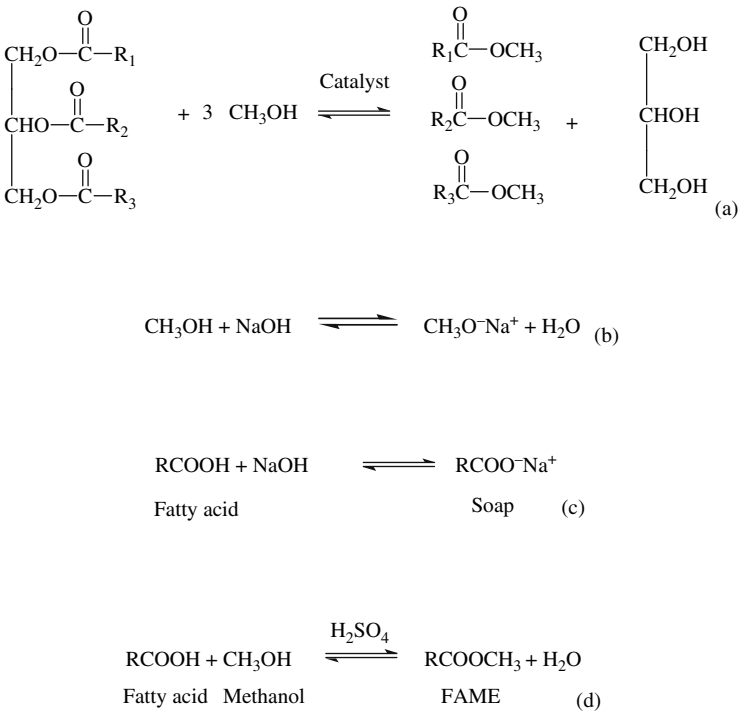


Figure 1. (a) Methanolysis of a generic triglyceride for the production of Fatty Acid Methyl Esters (FAME). (b) Formation of the methoxide. (c) Unwanted reaction of soap formation. (d) Acid catalysed esterification of free fatty acids

the presence of a high content of free fatty acids an acid-catalysed process is used (Fig. 1d). This employs strong acids such as H_2SO_4, as well as high temperatures and pressures, and requires longer reaction times than the alkaline process. Industrial plants where both alkaline and acid processes are performed require that reactors and accessories are resistant to these aggressive agents; moreover, high safety standards are needed. Due to these drawbacks, alternative and more environmentally sustainable routes for biodiesel production are being investigated.

2. BIODIESEL PRODUCTION USING LIPASES

Biodiesel production using lipases was first described by Mittlebach (1990) who showed that the lipase from *Pseudomonas fluorescens* was superior to those from *Candida* sp. and *Mucor miehei* for sunflower oil alcoholysis. The alcoholysis was carried out both in the presence of solvent (petroleum ether) and in solvent free conditions, and using five homologous alcohols with or without the addition of water. Since then, subsequent studies have focused on different lipases, different triglyceride feedstocks, different alcohols and different experimental conditions (temperature, water content, stoichiometric ratio between reagents, enzyme concentration, solvent use etc).

2.1. Lipases in Non-aqueous Media

It is now well established that enzymes can work with high activities in water-poor environments usually called non-conventional media (Vermue and Tramper, 1995; Salis *et al*, 2005a). The interest in using enzymes in non-aqueous media (organic solvents, supercritical fluids, solvent-free systems, gaseous media and ionic liquids) arises from the possibility to perform unusual reactions. Like other hydrolytic enzymes, lipases can function differently in such media, and instead of triglyceride hydrolysis they can catalyse transesterification reactions such as the alcoholysis involved in biodiesel production when suitable reagents and only limited amounts of water are present.

2.2. Sources of Lipases

Lipases used in biotechnology are normally of microbial origin (Jaeger and Eggert, 2002) and are produced by fermentative processes. A number of commercial lipases are available for applied biocatalysis (Pandey *et al*, 1999). Table 1 lists those most often utilised for biodiesel production. Whilst some are employed as free powders the majority are used as immobilized preparations. Some of the latter are commercially available, and in a number of cases the enzymes have been immobilized on different supports. References given in Table 1 cite first use of a lipase or its best performance.

Table 1. Source of free and immobilised lipases used for biodiesel production

Lipase source	Commercial name	Supplier	Support	Reference
Candida antarctica	SP435	Novo	Acrylic resin[(a)]	Nelson *et al.,* 1996
	Novozym 435	Novo	Acrylic resin[(a)]	Shimada *et al.,* 1999
	Chirazyme L-2	Roche	None	(Lee *et al.,* 2002)
Candida cylindracea	OF	Meito Sangyo	None	Lara and Park 2004
Candida rugosa	-	Meito Sangyo	None	Kaieda *et al.,* 2001
Chromobacterium viscoum	-	Asahi	Celite-545[(b)]	Shah *et al.,* 2004
Cryptococcus spp. S-2	Lipase produced in the researchers' laboratory		None	Kamini and Iefuji 2001
Porcine pancreatic	-	Sigma	Anion exchange resin[(a)]	Yesiloglu, 2004
Pseudomonas cepacia	PS	Amano	Sol-gel matrix[(b)]	Noureddini *et al.,* 2005
	PS	Amano	None	Kaieda *et al.,* 2001
	PS-30	Amano	None	Abigor *et al.,* 2000
	PS-30	Amano	Pyllosilicate sol-gel matrix[(b)]	Hsu *et al.,* 2002
	PS-D	Amano	Diatomaceous earth[(a)]	Salis *et al.,* 2005b
Pseudomonas fluorescens	-	Rhöm GmbH	None	Mittlebach, 1990
	AK	Amano	None	Kaieda *et al.,* 2001
	AK	Amano	Porous kaolinite[(b)]	Iso *et al.,* 2001
	AK	Amano	Polypropylene EP100[(b)]	Soumanou and Bornscheuer, 2003b
Mucor Miehei	Lipozyme IM60	Novo	Anion exchange resin[(a)]	Nelson *et al.,* 1996
Rhizopus oryzae	F-AP15	Amano	None	Kaieda *et al.,* 1999
Thermomyces lanuginosa	Lipozyme TL IM	Novo	Acrylic resin[(a)]	Du *et al.,* 2003
	-	Novo	Pyllosilicate sol-gel matrix[(b)]	Hsu *et al.,* 2004b

[(a)] : Commercially available immobilised lipases.
[(b)] : Lipases immobilised by researchers in their own laboratories.

2.3. Use of Immobilized Lipases

The use of immobilized enzymes confers two important advantages: i) the ability to recycle the catalyst and ii) the ability to perform continuous processes. Several reviews on this topic have been published (Fukuda *et al*, 2001; Shimada *et al*, 2002; Shah *et al*, 2003). A number of methods for the immobilisation of lipases on solid supports have been reported (Adlercreutz *et al*, 1996; Pedersen and Christensen, 2000). Among these, the best seem to be based on entrapment of the enzyme in hydrophobic sol-gel matrices (Reetz, 1997) or its adsorption onto hydrophobic supports such as polypropylene (Bosley and Peilow, 1997; Salis *et al*, 2003a). Commercially available lipases are supplied both as lyophilised powders, which contain other components in addition to the lipase (Salis *et al*, 2005d), and immobilied preparations. The immobilized lipase most frequently used for biodiesel production is lipase B from *Candida antarctica* (Nelson *et al*, 1996; Shimada *et al*., 1999; Samukawa *et al*, 2000; Watanabe *et al*, 2000; Watanabe *et al*, 2001; Bélafi-Bakó *et al*, 2002; Köse *et al*, 2002; Watanabe *et al*, 2002; Chen and Wu, 2003; De Oliveira *et al*, 2004; Du *et al*, 2004b; Tuter *et al*, 2004; Chang *et al*, 2005; Lai *et al*, 2005). This is supplied by Novozymes under the commercial name Novozym 435 (previously called SP435) and is immobilized on an acrylic resin. The *Mucor miehei* commercial lipase (Lipozyme IM60 - Novozymes) immobilized on a macroporous anionic exchange resin has also been extensively used for the same purpose (Mittlebach, 1990; Nelson *et al*., 1996; Selmi and Thomas, 1998; Dossat *et al*, 1999; Shieh *et al*, 2003; De Oliveira *et al*, 2004). Although commercially immobilized preparations may find immediate application, the development of new supports is of considerable interest.

Pseudomonas fluorescens lipase immobilized on porous kaolinite (Toyonite 200-M) gave high conversion ratios for propyl oleate and butyl oleate compared to those obtained with the lipases from *Pseudomonas cepacia, Mucor javanicus, Candida rugosa* and *Rhizopus niveus*. The *Pseudomonas cepacia* lipase (PS-30) immobilized on a phyllosilicate sol-gel matrix was found to be more active than the lipases of *Candida antarctica* and *Thermomyces lanuginosa* immobilized on granulated silica. It was suggested that the higher ester yields of lipase PS-30 may be due to entrapment of the lipase within the clay sol-gel matrix and its protection from methanol inactivation. Granulated lipase preparations do not protect the enzymes from inactivation by polar substrates (*i.e.* methanol) since they are adsorbed onto the support (Hsu *et al*, 2002). *Thermomyces lanuginosa* and *Pseudomonas cepacia* lipases immobilized on a phyllosilicate sol-gel matrix were shown to catalyse ester formation (80–90% yield) from greases containing a range of free fatty acids from 2.6 to 36% (Hsu *et al*, 2004b). *Porcine pancreatic* lipase immobilized by ionic linkage to a macroporous anion exchange resin was used for the ethanolysis of sunflower oil in a solvent-free system. High substrate conversion was obtained by performing the reaction with an oil:alcohol molar ratio of 1:3, at a temperature of 45 °C, 0% of added water and 10% wt of lipase based on the weight of the substrate (Yesiloglu, 2004). The choice of support seems to influence the methanolysis of

triolein in n-hexane. Although not described in detail, it has been reported that *Pseudomonas fluorescens* lipase was significantly more active when immobilized on polypropylene EP100 compared to celite. A conversion of 72.4% was achieved in the former case but only 1.5% in the latter (Soumanou and Bornscheuer, 2003b). It should be pointed out that these two supports have very different morphological features in terms of surface area, pore size distribution and chemical nature (Bosley and Peilow, 1997; Barros *et al*., 1998). These parameters strongly influence enzyme performance but this interesting subject has not been further investigated. Shah *et al*. (2004) immobilized *Chromobacterium viscosum* lipase on Celite 545 for the ethanolysis of *Jatropha* oil and found that this procedure increased ester yield from 62 %, obtained with the free lipase, to 71%.

A procedure for the immobilisation of *Pseudomonas cepacia* lipase was recently proposed by Noureddini *et al*. (2005). The lipase was gel-entrapped by polycondensation of hydrolysed tetramethoxysilane and *iso*-butyltrimethoxysilane. The immobilized lipase catalysed full triglyceride conversion in a very short time (30 min), it was very stable and lost little activity when subjected to repeated use.

3. SUBSTRATES USED FOR BIODIESEL PRODUCTION

3.1. Use of Different Oil/Fat Sources

The use of a triglyceride feedstock for biodiesel production depends on regional availability and economics. Rapeseed oil is the most widely used feedstock in Europe; soybean is mainly used in the United States, and palm oil is used in tropical regions (*i.e.* Malaysia). The main difference between these oils is their fatty acid composition, and this strongly affects some important features of the final biodiesel mixture. Table 2 details the compositions of the most common oils suitable for biodiesel production. The most abundant fatty acids are palmitic, stearic, oleic and linoleic acids. The main physical and chemical properties of an oil/fat depend on the chemical structures of its fatty acids (Fig. 2). In this regard, a frequent problem with biodiesel is its stability to oxidation. In linseed, sunflower and soybean oils the high contents of linoleic acid confers low stability to oxidation as a result of the presence of two double bonds. Indeed, the oxidation of unsaturated compounds

Table 2. Main fatty acids constituent of the most common oil/fat sources for biodiesel production

Oil/fat source	Palmitic acid (C16:0)	Stearic acid (C18:0)	Oleic acid (C18:1)	Linoleic acid (C18:2)	Problems
Soy bean oil	8	4	28	53	Oxidation stability
Palm oil	42	5	41	10	Low temperature
Rape seed oil	4	1	60	20	–
Sun flower oil	6	4	28	61	Oxidation stability
Beef tallow	26	18	37	10	Low temperature
Jatropha oil	13	7	45	34	Low temperature

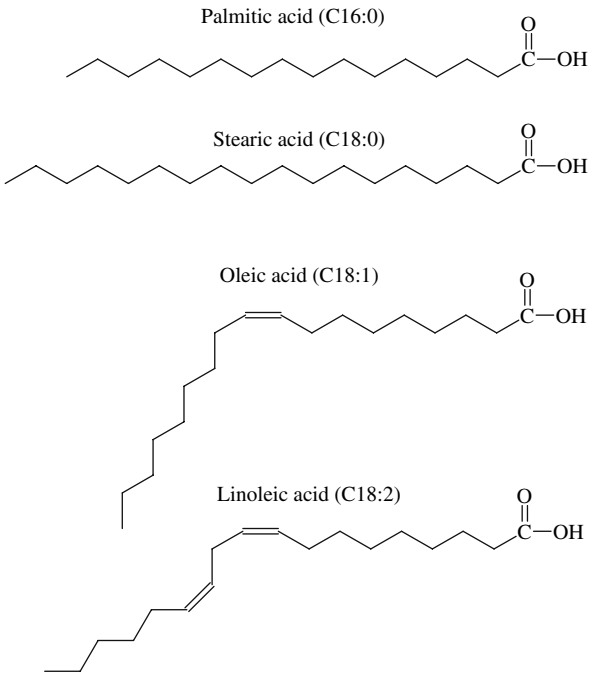


Figure 2. Chemical structure of the most common fatty acids occurring in oils and fats

proceeds at different rates depending on the numbers and positions of the double bonds. The CH_2 groups in allylic positions relative to the double bonds in the fatty acid chains are those susceptible to oxidation (Knothe, 2005b). By comparison, palm oil and animal fats contain high percentages of saturated fatty acids that are responsible for the poor low-temperature properties (*i.e.* high cloud point and pour point values) of biodiesel fuel. This constitutes a problem in cold regions during winter. From these considerations and the data in Table 2, it can be concluded that rapeseed oil is one of the most suitable sources for biodiesel production. Clearly, triglyceride source affects the properties of the biodiesel blends. Indeed, it was found that palm kernel oil ethyl esters have a viscosity of 9.33 mm^2/s, a cloud point of 12 °C and a pour point of 8 °C, whereas coconut 1-butyl esters have a viscosity of 7.34 mm^2/s, a cloud point of 5 °C and a pour point of −8 °C. The properties of these fuels are likely to be dependent on both the oil and the alcohol used in the transesterification (Abigor *et al*, 2000).

3.2. Vegetable Oils

The vegetable oils most studied for use in biodiesel production by biocatalysis (see Table 3) originate from: soy bean (Nelson *et al*, 1996; Kaieda *et al*, 1999; Kaieda *et al*, 2001; Shieh *et al*, 2003), sunflower (Mittlebach, 1990; Bélafi-Bakó

et al., 2002; Soumanou and Bornscheuer, 2003b) and rapeseed (Nelson *et al*., 1996). However, some other oleaginous, non-edible species should be mentioned: *Jatropha* (euphorbiaceae) is a plant that grows in harsh soils and its seed kernel is 40–60 % (w/w) oil. This plant cannot be used for edible purposes since its oil contains some toxic substances, *i.e.* phorbol esters, that render the oil unsuitable for use in cooking (Shah *et al*., 2004). Other triglyceride sources have been explored including the Nigerian lauric oils palm kernel oil and coconut oil (Abigor *et al*., 2000), rice bran oil (Kamini and Iefuji, 2001; Lai *et al*., 2005), refined cotton seed oil (Köse *et al*., 2002), peanut palmolein oil (Soumanou and Bornscheuer, 2003b) and castor oil (De Oliveira *et al*., 2004). Regarding biocatalytic processes, almost all sources of triglycerides can be considered equivalent as enzyme substrates. The different conversion percentages obtained from the transesterification of palm kernel oil and coconut oil with lipase PS30 are likely to be due to the different alcohols used (ethanol and butanol respectively) (Abigor *et al*., 2000).

3.3. Low Value Triglyceride Feedstocks

The main hurdle in the commercialisation of biodiesel is the cost of the raw material. Biodiesel – produced by base catalysis - cost more than 0.50 US$/dm^3 in 2001 as compared with 0.35 US$/dm^3 for petroleum-based diesel (Zhang *et al*., 2003). It has been reported that 60–75% of the price of biodiesel derives from the cost of the feedstock oil (Krawczyk, 1996). For this reason, low value triglyceride feedstocks are interesting alternatives for biodiesel production. The principle problem associated with their use is the necessity for preliminary treatments to render the oil/fat suitable for the transesterification process. Some of these can be performed by lipases. Attention has also been paid to the use of low-value triglycerides such as those from restaurant grease (Hsu *et al*., 2002), waste edible oil (Watanabe *et al*., 2001) and animal fats, *i.e.* tallow (Nelson *et al*., 1996).

Waste bleaching earths from crude vegetable oil refining processes contain approximately 40% oil by weight. Efficient methanolysis of oils recovered by organic solvent extraction - identified as originating from soybean, palm and rapeseed – has been reported for *Rhizopus oryzae* lipase in the presence of a high water content and a single injection of methanol (Lara and Park, 2003). In a follow-up study the same authors found that *Candida cylindracea* lipase was the most active enzyme in methanolysis of oil from waste activated bleaching earths when *n*-hexane was used as the solvent (Lara and Park, 2004).

Sunflower acid oils mainly consist of 55.6% free fatty acids and 24.7% triacylglycerols. They are the main by-product of the alkali refining process of crude vegetable oils to produce edible oils, and are obtained by acidification of soapstocks. This waste oil was transformed into FAME (65% yield) by means of immobilized *Candida antarctica* lipase B (15% based on acid oil weight) at 40 °C after 1.5 h and using *n*-hexane as the solvent (Tuter *et al*., 2004).

As already mentioned, animal fat produces a biodiesel with poor low-temperature properties. In order to improve cold temperature resistant biodiesel several strategies

Table 3. Biodiesel production through different triglyceride feedstocks, alcohols, solvents, reactor types

Oil/fat source	Alcohol	Lipase source	Solvent	Type of reactor	Conversion (c) or yield (y) (mol or wt%)	Reference
Sunflower oil	Ethanol	*Pseudomonas fluorescens*	Petroleum ether	Batch	82 (y)	Mittlebach, 1990
	Methanol	*Pseudomonas fluorescens*	Solvent free	3-step batch	> 90	Soumanou and Bornscheuer, 2003b
	Methanol	*Candida antarctica*	Solvent free	Membrane reactor	97 (c)	Bélafi-Bakó *et al.,* 2002
	Ethanol	*Mucor miehei*	Solvent free	Batch (4 cycles)	83 (y)	Selmi and Thomas, 1998
	Ethanol	*Porcine pancreatic*	Solvent free	Batch	81 (y)	Yesiloglu, 2004
	Methanol	*Rhizomucor miehei*	Solvent free	3-step batch (8 cycles)	> 80 (c)	Soumanou and Bornscheuer, 2003a
High oleic sunflower oil	Butanol	*Rhizomucor miehei*	*n*-Hexane	Packed bed reactor	> 80 (c)	Dossat *et al.,* 1999
Sunflower acid oil	Methanol	*Candida antarctica*	*n*-Hexane	Batch	63.6 (y)	Tuter *et al.,* 2004
Soybean oil	Methanol	*Pseudomonas cepacia*	Solvent free	Batch	~ 60(y)	Kaieda *et al.,* 2001
	Methanol	*Candida antarctica*	Solvent free	3-step batch (20 cycles)	97 (y)	Samukawa *et al.,* 2000
	Methanol	*Thermomyces lanuginosa*	Solvent free	Continuous batch	80–90 (y)	Du *et al.,* 2003
	Methanol	*Thermomyces lanuginosa*	Solvent free	3-step batch (15 cycles)	94 (y)	Xu *et al.,* 2004
	Methanol – ethanol	*Pseudomonas cepacia*	Solvent free	Batch (12 cycles)	67 (y)	Noureddini *et al.,* 2005

(Continued)

Table 3. (Continued)

Oil/fat source	Alcohol	Lipase source	Solvent	Type of reactor	Conversion (c) or yield (y) (mol or wt%)	Reference
	Methanol	*Rhizopus oryzae*	Solvent free	Batch	80 (y)	Kaieda *et al.*, 1999
Degummed soybean oil	Methanol	*Candida antarctica*	Solvent free	3-step batch (25 cycles)	93.8 (c)	Watanabe *et al.*, 2002
Soybean and rapeseed oil mixture	Methanol	*Candida antarctica*	Solvent free	3-step batch 50 cycles	98.4 (c)	Shimada *et al.*, 1999
	Methanol	*Candida antarctica*	Solvent free	3-packed-bed reactors (100 days)	93 (y)	Watanabe *et al.*, 2000
Triolein – safflower oil	1-propanol	*Pseudomonas fluorescens*	1,4-dioxane	Batch		Iso *et al.*, 2001
	1-propanol		Solvent free	Batch (10 cycles)		
Triolein	Fusel oil-like mixture	*Pseudomonas cepacia*	Solvent free	Batch	100 (c)	Salis *et al.*, 2005b
Nigerian lauric oils (palm kernel and coconut)	Ethanol	*Pseudomonas cepacia*	Solvent free	Batch	72 (c)	Abigor *et al.*, 2000
	1-butanol	*Pseudomonas cepacia*	Solvent free	Batch	40 (c)	
Castor oil	Ethanol	*Rhizomucor miehei*	*n*-Hexane	Batch	98 (c)	De Oliveira *et al.*, 2004
Cotton seed-oil	Primary and secondary alcohols	*Candida antarctica*	Solvent free	Batch	91.5 (methanol)	Köse *et al.*, 2002
	Methanol	*Rhizomucor miehei*	Solvent free	Batch (8 cycles)	> 90 (c)	Soumanou and Bornscheuer, 2003b

Rice bran oil	Methanol	*Cryptococcus* spp. S-2	Solvent free	Batch	80.2 (y)	Kamini and Iefuji, 2001
	Methanol	*Candida antarctica*	Solvent free	Batch	98 (c)	Lai *et al.*, 2005
Jatropha oil	Ethanol	*Chromobacterium viscoum*	Solvent free	Batch	92 (y)	Shah *et al.*, 2004
Waste activated bleaching earths (ABE) oil	Methanol	*Candida cylindracea*	Diesel fuel	Batch	~ 100 (y)	Kojima *et al.*, 2004
	Methanol	*Rhizopus oryzae*	Water	Batch	55 (y)	Lara Pizarro and Park, 2003
Waste edible-oil	Methanol	*Candida antarctica*	Solvent free	Packed-bed reactor (100 days)	90 (y)	Watanabe *et al.*, 2001
Tallow (other oils)	Primary alcohols	*Mucor Miehei*	*n*-Hexane	Batch	> 90 (c)	Nelson *et al.*, 1996
Restaurant grease	Methanol	*Pseudomonas cepacia*	Solvent free	Batch	98 (y)	Hsu *et al.*, 2002
	Methanol	*Candida antarctica*	Solvent free	Batch	96 (c)	Lee *et al.*, 2002
	Ethanol	*Burkholderia cepacia*	Solvent free	Packed bed reactor	> 96 (y)	Hsu *et al.*, 2004a
Fractionated lard	Methanol	*Candida antarctica*	Solvent free	Batch	58 (c)	Lee *et al.*, 2002

can be followed. Lee *et al.* (2002) decreased the content of saturated fatty acids present in lard and restaurant grease by performing an acetone fractionation step followed by methanolysis catalysed by Chirazyme L-2 (*Candida antarctica* lipase).

Methanolysis of rice bran oil having a free fatty acid content greater than 18% gave conversions < 68%. A two-step lipase-catalysed (*Candida antarctica*) methanolysis of rice bran oil was developed for the efficient conversion of both free fatty acids and acylglycerides to FAME. More than 98% conversion can be obtained in 4–6 h, depending on the relative proportion of free fatty acids and acylglycerides present (Lai *et al.*, 2005).

3.4. Alcohols

As already discussed, for cost reasons methanol is the reagent most frequently used for triglyceride transesterification. Nevertheless, other alcohols are also used. In Brazil biodiesel is produced by ethanolysis of triglycerides since ethanol is obtained cheaply by the fermentation of sucrose from sugarcane. The use of different alcohols gives different results. Alcoholysis of Nigerian lauric oils catalysed by lipase PS-30 gave different oil conversions with methanol, ethanol, 1-propanol, iso-propanol, 1-butanol and iso-butanol (Abigor *et al.*, 2000). However, this was not only related to the alcohol since the conversion trend was different for palm kernel compared to coconut oil. It is worth noting that methanol gave the lowest conversions in both these cases.

Nelson *et al.* (1996) used linear and branched alcohols for the biocatalytic transesterification of tallow using hexane as solvent. They found that *Candida antarctica* lipase was the most efficient in the transesterification with secondary alcohols, whereas lipase from *Mucor miehei* was the most efficient with primary alcohols.

The use of C_3-C_5 linear and branched alcohols from fusel oil, a low-value residue from ethanol distillation, might constitute an interesting and cheap alternative to methanol. Salis *et al.* (2005b) carried out the biocatalytic alcoholysis of triolein with a fusel-oil like mixture. On a molar basis, fusel oil mainly comprises: isoamyl alcohol (64.4%), 2-butanol (27.6%), 2-methyl-1-propanol (12.3%), 1-propanol (5.6%) and 1-butanol (1.3%). These alcohols are not enzyme denaturing and their esters, mainly the branched ones, improve the low-temperature properties of biodiesel blends (Dunn, 2005). It should be remarked that the absence of methanol makes the whole process more environmentally friendly. A different result was obtained by Kaieda *et al.* (1999). In their case the ester content decreased in the series methanol > iso-butanol > ethanol > butanol > propanol in catalytic alcoholysis of rice bran oil using crude *Cryptococcus* spp. S-2 lipase. A methyl ester content of 80.2% was obtained in the presence of a high content of water (80% of substrate weight).

Conversion of cottonseed oil in a 24 h reaction at 40 °C in a solvent free system has been performed with various alcohols. Low conversions (10%) were observed with short chain alcohols especially with *Thermomyces lanuginosus* lipase. Higher conversions were obtained with *Rhizomucor miehei* lipase (about 30%)

and *Pseudomonas fluorescens* lipase (about 57%). Using secondary alcohols as substrates for ester production, conversion by 2-propanol was found to be lower than that by isobutanol. In all cases, better conversion levels were obtained when cottonseed oil was transesterified by 1-propanol and isobutanol (Soumanou and Bornscheuer, 2003b).

4. REACTION CONDITIONS FOR BIODIESEL PRODUCTION

4.1. Solvent Versus Solvent-free Processes

Miscibility between methanol and triglycerides is very poor, thus the presence of a solvent helps to form a monophasic system. However, the use of a solvent has some disadvantages such as its storage before use and its removal and disposal after the process both of which generate environmental problems. Kamini and Iefuji (2001) found that addition of organic solvents (10% w/v) increased the content of methyl esters from 80.2% (no solvent) to 85% (DMSO), 86.1% (*n*-hexane) and 87.2% (petroleum ether), whereas the addition of diethylether decreased methyl ester content to 70.4%. The same authors stated that although the addition of solvents gives favourable results in some cases, their use is not recommended for reasons of economy and safety. The same conclusions were arrived at by Iso *et al.* (2001). They reported that the methanolysis and ethanolysis of triolein and safflower oil were better carried out in the presence of 1,4-dioxane as solvent. Conversely, activity in other solvents such as benzene, chloroform and tetrahydrofuran was extremely low. As an alternative they proposed the use of 1-propanol and 1-butanol which allow full miscibility with the oil and yield high reaction rates.

Soumanou and Bornscheuer (2003b) found that *n*-hexane was necessary to perform methanolysis of various vegetable oils. The highest conversion (97%) was obtained with *Thermomyces lanuginosa* lipase after 24 h. By contrast, this lipase was found to be almost inactive in solvent-free reaction medium where methanol or 2-propanol were the alcohol substrates. In successive work these authors investigated the effect of other solvents, namely cyclohexane, *n*-heptane, isooctane, acetone and petroleum ether. All gave conversions in the range 60–80% with three immobilized lipases (from *Mucor miehei, Thermomyces lanuginosa* and *Pseudomonas fluorescens*). Acetone, by comparison, gave a conversion in the range 5–20%. This polar solvent may alter the native conformation of the enzyme by disrupting hydrogen bonding and hydrophobic interactions, thereby leading to a very low alcoholysis rate (Soumanou and Bornscheuer, 2003b).

An interesting solvent process was proposed by Kojima *et al.* (2004). The enzymatic methanolysis of waste oil from activated bleaching earth (ABE) was performed by using diesel fuel as solvent (0.6 mL/g of waste ABE). In this system, the lipase from *Candida cylindracea* showed high stability and activity reaching approximately 100% yield of FAME in 3 h. Fuel analysis showed that the FAME produced in this way complied with the Japanese diesel standard. The use of diesel oil as solvent in FAME production from the waste ABE simplified the process since there is no need to separate the organic solvent from the FAME-solvent mixture.

4.2. Water Content

Water content in biocatalysis in non-conventional media is a very important parameter since it strongly affects enzymatic activity (Adlercreutz, 2000; Salis *et al*, 2005a). The reaction rates of methanolysis catalysed by the *Candida rugosa* and *Pseudomonas fluorescens* lipases decreased significantly when water content was low since water prevents the inactivation of these lipases by methanol (Kaieda *et al*, 2001). Conversely, the rate of the same reaction catalysed by *Pseudomonas cepacia* lipase remained high even under low water content conditions. Shah *et al* (2004) found that the addition of water to *Chromobacterium viscosum* lipase increased ethyl ester yield from 62 to 73% for the free preparation (1% of water) and from 71 to 92% for the immobilized (on celite) preparation (0.5% of water). *Cryptococcus* spp. S-2 lipase was able to catalyse methanolysis of rice bran oil in the presence of high water content (80% by weight of substrate), thus obtaining a methyl ester yield of 80.2% after 120 h (Kamini and Iefuji, 2001).

There are conflicting reports on the effect of water content in the synthesis of biodiesel. Some authors state that ester yield increases if high amounts of water are present in the system (Kamini and Iefuji, 2001), whilst others claim the contrary (Hsu *et al*, 2002). This apparent contradiction might be explained by considering that the effect of water in these systems depends on the enzyme, the support and the medium (solvent or solvent-free). As a general consideration, high water content should decrease ester yield since undesirable triglyceride hydrolysis occurs. Although water is involved in several enzyme denaturing processes, a certain amount of water must be present to 'lubricate' polypeptide chains and keep an enzyme in its active conformation. In this regard, water content is better expressed in terms of water activity (Halling, 1994) since only the use of this thermodynamic parameter allows a real comparison of the effects of water among different enzymes in the same medium. Unfortunately, the majority of papers reporting studies on the effect of water cite water concentration instead of water activity.

4.3. Biocatalyst Recycling and Continuous Processes

One of the main drawbacks to the biocatalytic production of biodiesel is the cost of the enzyme. Enzyme recycling might be the solution to this problem.

Pseudomonas fluorescens lipase immobilized on kaolinite lost one third of its activity when used for the second time, but no further decrease was observed in successive applications. The initial decrease in activity was put down to enzyme desorption from the solid support that was not observed after repeated (10 times) use (Iso *et al*, 2001). Repeated batch reactions showed that *Mucor miehei* lipase showed high stability, retaining about 70% of its initial conversion after 8 cycles (24 h each cycle), whereas *Thermomyces lanuginosa* retained only 35% of the initial conversion under the same experimental conditions. This difference was ascribed to various factors such as inactivation of the biocatalyst in the oil phase, the type of carrier used for the immobilisation or enzyme sensitivity to long-term methanol exposure (Soumanou and Bornscheuer, 2003b).

A two-step batch process employing Novozym 435 was proposed in which 1/3 equivalent of methanol is added initially and the remaining 2/3 is added in the second step. More than 95 % conversion was observed. When the reaction was repeated by transferring the immobilized lipase to a fresh substrate mixture, it was observed that the enzyme could be reused for 70 cycles (105 days) without any decrease in the conversion. However, when the two-step reaction was conducted using a reactor equipped with an impeller, the support was destroyed thus limiting reusability to just a few cycles (Watanabe *et al*., 2000). As an alternative to recycling the biocatalyst, continuous processes might help to overcome the economic drawbacks of enzymatic biodiesel production.

Du *et al*. (2003) studied the effects of temperature, the oil/alcohol molar ratio and the presence of the by-product glycerol during Lipozyme TL IM-catalysed continuous batch operation when short-chain alcohols were used as the acyl acceptor. In non-continuous batch operation the oil/alcohol ratio and temperature optima were 1:4 and 40–50 °C, respectively; however, during continuous batch operation the optimal conditions were found to be 1:1 and 30 °C. Enzymatic activity was noted to be 95% after 10 cycles when iso-propanol was used to remove glycerol during repeated use of the lipase. It was also found that the enzyme lost its activity rapidly at 40 °C. 30 °C was found to be the optimal temperature at which lipase retains high activity, even after many batches. Continuous and discontinuous processes have different optimal temperatures. Presumably, lipases exhibit relatively high activity during short-term operation in non-continuous reactions, but during long-term operation they may lose their activity rapidly, particularly at relatively high temperatures.

The lipase from *Burkholderia cepacia*, immobilized on a phyllosilicate sol-gel matrix, was used for the continuous production of ethyl esters of grease (Hsu *et al*., 2004a). An ester yield>96% was obtained.

4.4. Lipase Inhibition and Regeneration

Both reagents and products of the alcoholysis reaction and some substances contained in the feedstock can inhibit lipases. Below are described the different kinds of inhibition and their possible remediation. A schematic summary is also shown in Table 4.

4.4.1. Methanol inhibition

Methanol is poorly miscible with oil/fat and tends to inactivate enzymes. *Candida antarctica* lipase B, an enzyme that is very stable in organic media (Salis *et al*., 2003b), was found to be progressively inactivated by methanol in amounts above 1/2 molar equivalent. To overcome this problem, multi-step addition of methanol was proposed (Shimada *et al*., 1999): only 1/3 of the stoichiometric amount is used in the first step. After this first addition, a mixture of methanol, FAME, mono- and di-glycerides is formed. The system thus composed is less aggressive toward the enzyme compared to the initial conditions when only the oil and alcohol are present (Shimada *et al*., 2002). It has also been shown that pre-incubation of Novozym 435

Table 4. Type of inhibition and possible remediation for discontinous and continuous enzymatic production of biodiesel.

Type of process	Lipase	Inhibitor	Remediation	Reference
Discontinuous (Batch)	*Candida antarctica*	Methanol	Three step addition of methanol	Shimada *et al.,* 1999
	Candida antarctica	Methanol	Addition of 10 wt% of silica gel	Lee *et al.,* 2002
	Candida antarctica	Methanol	Lipase immersion pre-treatment in 2-butanol or *tert*-butanol	Chen and Wu, 2003
	Candida antarctica	Phospholipids	Oil degumming	Watanabe *et al.,* 2002
	Candida antarctica	Phospholipids	Lipase immersion pre-treatment	Du *et al.,* 2004b
	Candida antarctica	Phospholipids	Simultaneous dewaxing and degumming	Lai *et al.,* 2005
	Thermomyces lanuginosa	Methanol	Three step addition of methanol	Xu *et al.,* 2004
		Glycerol	Enzyme washing with isopropanol	
Continuous (Packed-bed reactor)	*Rizhomucor miehei*	Glycerol	Addition of silica gel; Use of hexane amended with acetone; Rinsing of the catalyst	Dossat *et al.,* 1999
	Candida antarctica	Methanol	Three-step flow reaction	Watanabe *et al.,* 2002
	Candida antarctica	Glycerol	Use of a dialysis membrane	Bélafi-Bakó *et al.,* 2002

in methyl oleate for 0.5 h followed by soybean oil for 12 h resulted in a methyl ester yield of 97% within 3.5 h (Samukawa *et al*., 2000).

Chirazyme L-2 (*Candida antarctica* lipase B) activity was found to be inhibited in the presence of more than one mol of methanol. Therefore methanol was added to the reaction in three consecutive steps. It was also found that addition of 10% silica gel to the reaction mixture increased the conversion rate. It has been postulated that the excess methanol forms a barrier around the lipase structure that may involve the active site, thereby hindering contact with the acyl donor and leading to inactivation of the lipase. Silica gel may act as a methanol depot thus preventing direct exposure of the lipase to high concentrations of methanol and resulting in an increased reaction rate (Lee *et al*., 2002).

Chen (2003) proposed a procedure for enzyme regeneration that solves the problem caused by methanol. They found that an alcohol with three or more carbon atoms, preferably 2-butanol or *tert*-butanol, can regenerate deactivated immobilized enzyme. The procedure consists of an immersion pre-treatment of the enzyme in the alcohol. By this means the activity of Novozym 435 was increased about ten-fold compared to the untreated enzyme. Following complete deactivation of the enzyme by methanol, washing with 2-butanol or *tert*-butanol successfully regenerated the enzyme restoring about 56 % and 75 % of its original activity, respectively.

Another way to overcome the methanol inhibition of *Candida antarctica* lipase was developed by Xu *et al*. (2003). They proposed the use of methyl acetate as a novel acyl acceptor for biodiesel production. When used for soybean oil transesterification it gave 92% methyl ester yield with a molar ratio methyl acetate:oil = 12:1. Unlike the case for methanol (inhibition at ratios greater than 1:1) no adverse effects were observed. In addition, the use of methyl acetate enabled crude soybean oil to be used, attaining the same yield as in the case of refined oil. No activity loss was observed with methyl acetate as the acyl acceptor upon repeated use (100 batches) of the lipase. Moreover, the by-product triacetylglycerol is an important chemical of greater economic value than glycerol (Du *et al*., 2004a). The kinetics of the lipase catalysed inter-esterification of triglycerides with methyl acetate was studied by Xu *et al*. (2005). Three consecutive and reversible reactions occurred in the inter-esterification. The main finding was that the first reaction step was limiting for the overall process.

Distinctly from other microbial lipases, *Pseudomonas cepacia* lipase was found to give high methyl ester contents in reaction mixtures containing methanol:oil equivalent ratios of up to 2–3:1. This enzyme seems to be substantially resistant to methanol (Kaieda *et al*., 2001). Also *Rhizopus oryzae* lipase was not inactivated by methanol (Lara and Park, 2003).

4.4.2. Glycerol inhibition

Glycerol has a negative effect on FAME production by *Thermomyces lanuginosa* lipase (Selmi and Thomas, 1998). Enzyme washing with iso-propanol was found to be effective at removing glycerol. In the presence of iso-propanol the lipase showed

relatively high activity and a methyl ester yield of over 94% was retained after repeated use over 15 batches (Xu *et al*., 2004).

4.4.3. Inhibition by phospholipids

Watanabe *et al*. (2002) found that gum present in crude soybean oil inhibits *Candida antarctica* lipase since this negative effect was not observed with degummed oil. The main components of gum are phospholipids (PLs). The inhibition observed may be due to PLs bound to the immobilized preparation that interfere in the lipase-substrate interaction. When degummed oil was used it was successfully converted (93.8%) into its corresponding methyl esters, and the lipase could be reused for 25 cycles without any loss of activity. Similarly, Du *et al*. (2004b) found that the methyl ester yield was significantly lower with crude compared to refined soya-bean oil. The major difference between refined and crude oils was found to be in their contents of PLs, free acid and water, which have various influences on biodiesel production. PL content was found to be the most important parameter: the higher the PL content in the oil, the lower the methyl ester yield. Water and free fatty acid content did not affect ester yield to the same extent as PL content. The inhibitory effect of PLs was overcome by an immersion pre-treatment of the lipase (from *Candida antarctica*) in crude oil for 120 h. This resulted in a methyl ester yield of 94% that obtained with refined oils (Du *et al*., 2004b). Lai *et al*. (2005) observed inactivation of lipase by PLs and other minor components during the methanolysis of crude rice bran oil. To avoid this, simultaneous dewaxing/degumming was found to be efficient.

4.4.4. Inhibition during continuous operation

Glycerol is formed during the alcoholysis of triglycerides. If this reaction is performed in a Packed-bed reactor glycerol tends to adsorb onto the hydrophilic supports instead of being withdrawn by the flow. This adsorption inhibits enzyme activity. Two possible mechanisms of inhibition have been proposed: the adsorption of glycerol may cause a decrease in the (thermodynamic) water activity of the enzyme, or glycerol may form a layer around the enzyme that inhibits diffusion of the hydrophobic substrate. Studies on the butanolysis of high oleic sunflower oil (> 80% oleic acid content) catalysed by Lipozyme (the lipase from *Rhizomucor miehei*) favoured the second hypothesis. Various procedures were tested for their ability to retain the high initial enzymatic activity: addition of silica gel to the reactor bed in order to address the preferential adsorption of the glycerol; use of *n*-hexane amended with acetone as a reaction medium; and a semi-continuous process consisting of a transesterification reaction and rinsing of the catalyst in order to evacuate the adsorbed glycerol out of the reactor. This latter procedure permitted the restoration of initial enzymatic activity by using a rinsing solution amended with water with the thermodynamic water activity adjusted to an optimal value (Dossat *et al*., 1999).

When immobilized *Candida antarctica* lipase was used for the continuous methanolysis of a mixture of soybean and rapeseed oils in a Packed-bed reactor,

two kinds of enzyme inhibition were identified (Watanabe *et al*, 2000). One was due to the substrate (methanol) and the other to the by-product (glycerol) (Bélafi-Bakó *et al*, 2002). The first problem was overcome by performing a three-step flow reaction. As regards the second, it was proposed that glycerol inhibition could be eliminated by continuous operation in a membrane bioreactor using a suitable dialysis membrane. The reaction takes place in the primary side of the module, and the glycerol produced during the reaction passes through the membrane and is accumulated in the secondary (aqueous phase) side (Bélafi-Bakó *et al*, 2002).

4.5. Statistical Optimisation of Reaction Parameters for Biodiesel Production

The parameters – temperature, reaction time, amount of lipase, mole ratio of reactants – of the enzymatic (*Pseudomonas cepacia* lipase) ethanolysis of restaurant grease were optimised by response surface methodology (RSM). The regression equation obtained by RSM predicted optimal reaction conditions under which the predicted yield of ethyl ester was 85.4%. Subsequent experiments, using combinations of the theoretical parameters, revealed a trend in which experimental percentage yields of ethyl esters were consistently lower than the predicted values. This problem was overcome by adding 5% SP435 (*Candida antarctica* lipase) one hour after the beginning of the reaction. This resulted in an ester yield greater than 96% (Wu *et al*, 1999).

The ability of commercial immobilized *Mucor miehei* lipase (Lipozyme IM-77) to catalyse the methanolysis of soybean oil was investigated. RSM and 5-level-5-factor central composite rotatable design were used to evaluate the effects on reaction time, temperature, enzyme amount, molar ratio between reagents, and added water content on the percentage weight conversion to soybean methyl ester by transesterification. Experiments performed under these experimental conditions gave a molar conversion of 90.9% that should be compared with the theoretical value of 92.2% (Shieh *et al*, 2003). The same method was used to optimise the methanolysis of canola oil using Novozym 435 as catalyst (Chang *et al*, 2005). As in previous cases, the investigated parameters were reaction time, temperature, enzyme concentration, substrate molar ratio, and added water. Reaction temperature and enzyme concentration were the most important variables. High temperature and excess methanol inhibited the ability of Novozym 435 to catalyse the synthesis of biodiesel. Based on analysis of ridge max, the optimum synthesis conditions predicted a weight conversion value of 99.4%, very close to the actual experimental value (97.9% weight conversion).

De Oliveira *et al*. (2004) adopted a ‘Taguchi experimental design’ considering the following variables: temperature (35–65 °C), water (0–10 wt%), enzyme (5–20 wt%) concentrations, and oil to alcohol molar ratio (1:3 to 1:10) for the enzymatic ethanolysis of castor oil. Novozym 435 and Lipozyme IM were the catalysts.

Experimental conversions of 81.4% and 98% were observed for the two enzymes respectively, in agreement with the theoretical values (82% and 99.6%).

5. CONCLUSIONS

Lipases are suitable catalysts for biodiesel production. They catalyse both the alcoholysis of triglycerides and the esterification of free fatty acids. This is especially important for low value oil/fat of high acidity. Indeed, such substrates would otherwise require pre-treatment or an esterification step in the presence of an acid catalyst followed by the alkaline transesterification. Lipases can catalyse both reactions in the same reactor. These enzymes can work in the presence or in the absence of a solvent, but the second option is preferred in order to avoid the environmental and safety problems associated with the use of solvents. Various substances can inhibit lipase activity (methanol, glycerol, phosphlipids); however several ways have been proposed to resolve these inconveniences. Lipases work under mild operating conditions at room temperature and at atmospheric pressure. Consequently, energy-saving processes can be carried out. The biocatalytic process does not produce soaps or other by-products. If the reaction reaches completeness, only esters and glycerol are produced. This simplifies purification steps and consequentially reduces plant costs. A possible flow diagram for biodiesel production by lipases is outlined in Fig. 3. Once immobilized, lipases can be used several times or even in continuous processes. This overcomes the main disadvantage related to their high cost. In conclusion, biodiesel is an environmentally friendly fuel which, if obtained by biocatalysis, will contribute to reducing negative impacts on the environment. It may thus be justifiable to rename it 'bio-biodiesel'.

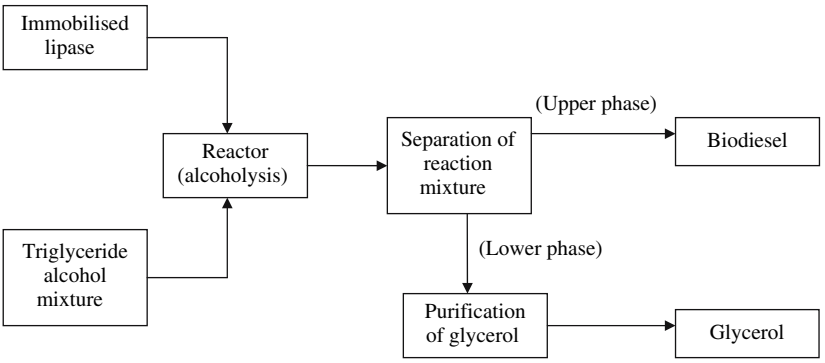


Figure 3. Flow diagram of biodiesel production through lipases

ACKNOWLEDGEMENTS

MIUR and CSGI (Firenze) are acknowledged for their financial support.

REFERENCES

Abigor, R.D., P.O. Uaudia, T.A. Foglia, M.J. Hass, K.C. Jones, E. Okpefa, J.U. Obibuzor and M.E. Bafor (2000). Lipase-catalysed production of biodiesel fuel from some Nigerian lauric oils. Biochemical Society Transactions *28*, 979–981.

Adlercreutz, P. (2000). Biocatalysis in non-conventional media. In: Applied Biocatalysis, A.J.J. Straathof and P. Adlercreutz. (eds.), Harwood Academic Publishers, pp. 295–316.

Adlercreutz, P., R. Barros and E. Wehtje (1996). Immobilisation of enzymes for use in organic media. Annals NY Accad. Sci. *795*, 197–200.

Barros, R.J., E. Wehtje, F.A.P. Garcia and P. Adlercreutz (1998). Physical characterization of porous materials and correlation with the activity of immobilized enzyme in organic medium. Biocatal. Biotransf. *16*, 67–85.

Bélafi-Bakó, K., F. Kovács, L. Gubicza and J. Hancsók (2002). Enzymatic biodiesel production from sunflower oil by *Candida antarctica* lipase in a solvent-free system. Biocatal. Biotransf. *20*, 437–439.

Bosley, J.A. and A.D. Peilow (1997). Immobilization of lipases on porous polypropylene: reduction in esterification efficiency at low loading. J. Am. Oil Chem. Soc. *74*, 107–111.

Chang, H.-M., H.-F. Liao, C.-C. Lee and C.-J. Shieh (2005). Optimized synthesis of lipase-catalyzed biodiesel by Novozym 435. J. Chem. Technol. Biotechnol. *80*, 307–312.

Chen, J.-W. and W.-T. Wu (2003). Regeneration of immobilized *Candida antarctica* lipase for transesterification. J. Biosci. Bioeng. *95*, 466–469.

De Oliveira, D., M. Di Luccio, C. Faccio, C. Dalla Rosa, J.P. Bender, N. Lipke, S. Menoncin, C. Amroginski and J.V. De Oliveira (2004). Optimization of enzymatic production of biodiesel from castor oil in organic solvent medium. Appl. Biochem. Biotechnol. *113–116*, 771–780.

Dossat, V., D. Combes and A. Marty (1999). Continuous enzymatic transesterification of high oleic sunflower oil in a packed-bed reactor: influence of the glycerol production. Enzyme Microb. Technol. *25*, 194–200.

Du, W., Y. Xu and D. Liu (2003). Lipase-catalysed transesterification of soya bean oil for biodiesel production during continuous batch operation. Biotechnol. Appl. Biochem. *38*, 103–106.

Du, W., Y. Xu, D. Liu and J. Zeng (2004a). Comparative study on lipase-catalyzed transformation of soybean oil for biodiesel production with different acyl acceptors. J. Mol. Catal. B: Enzym. *30*, 125–129.

Du, W., Y. Xu, J. Zeng and D. Liu (2004b). Novozym 435-catalysed transesterification of crude soya bean oil for biodiesel production in a solvent-free medium. Biotechnol. Appl. Biochem. *40*, 187–190.

Dunn, R.O. (2005). Cold weather properties and performance of biodiesel. In: The Biodiesel Handbook, G. Knothe, J. Van Gerpen and J. Krahl (eds.), AOCS Press, pp. 83–121.

Fukuda, H., A. Kondo and H. Noda (2001). Biodiesel fuel production by transesterification of oil. J. Biosci. Bioeng. *92*(5), 405–416.

Halling, P.J. (1994). Thermodynamic predictions for biocatalysis in non-conventional media: theory, tests, and recommendations for experimental design and analysis. Enzyme Microb. Technol. *16*, 178–206.

Hsu, A.-F., K. Jones, T.A. Foglia and W.N. Marmer (2002). Immobilized lipase-catalysed production of alkyl esters of restaurant grease as biodiesel. Biotechnol. Appl. Biochem. *36*, 181–186.

Hsu, A.-F., K.C. Jones, T.A. Foglia and W.N. Marmer (2004a). Continuous production of ethyl esters of grease using an immobilized lipase. J. Am. Oil Chem. Soc. *81*, 749–752.

Hsu, A.-F., K.C. Jones, T.A. Foglia and W.N. Marmer (2004b). Transesterification activity of lipases immobilized in a phyllosilicate sol-gel matrix. Biotechnol. lett. *26*, 917–921.

Iso, M., B. Chen, M. Eguchi, T. Kudo and S. Shrestha (2001). Production of biodiesel fuel from triglycerides and alcohol using immobilized lipase. J. Mol. Catal. B: Enzym. 16, 53–58.

Jaeger, K.-E. and T. Eggert (2002). Lipases for biotecnology. Curr. Opin. Biotechnol. *13*, 390–397.

Kaieda, M., T. Samukawa, T. Matsumoto, K. Ban, A. Kondo, Y. Shimada, H. Noda, F. Nomoto, K. Ohtsuka, E. Izumoto and H. Fukuda (1999). Biodiesel fuel production from plant oil catalyzed by rhizopus oryzae lipase in a water-containing system without an organic solvent. J. Biosci. Bioeng. *88*, 627–631.

Kaieda, M., T. Samukawa, A. Kondo and H. Fukuda (2001). Effect of methanol and water contents on production of biodiesel fuel from plant oil catalyzed by various lipases in a solvent-free system. J. Biosci. Bioeng. *91*, 12–15.

Kamini, N.R. and H. Iefuji (2001). Lipase catalyzed methanolysis of vegetable oils in aqueous medium by *Cryptococcus* spp. S-2. Process Biochem *37*, 405–410.

Knothe, G., Ed. (2005a). The Biodiesel Handbook. Champaign, AOCS Press.

Knothe, G. (2005b). Oxidative stability of Biodiesel. The Biodiesel Handbook. G. Knothe, J. Krahl and J. Van Gerpen (eds), Champaign, AOCS Press, pp. 122–126.

Kojima, S., D. Du, M. Sato and E.Y. Park (2004). Efficient production of fatty acid methyl ester from waste activated bleaching earth using diesel oil as organic solvent. J. Biosci. Bioeng. *98*, 420–424.

Köse, O., M. Tüter and H.A. Aksoy (2002). Immobilized Candida antarctica lipase-catalyzed alcoholysis of cotton seed oil in a sovent-free medium. Biores. Technol. *83*, 125–129.

Krawczyk, T. (1996). Biodiesel-alternative fuel makes inroads but hurdles remains. Inform *7*, 801–829.

Lai, C.-C., S. Zullaikah, R.S. Vali and Y.-H. Ju (2005). Lipase catalyzed production of biodiesel from rice bran oil. J. Chem. Technol. Biotechnol. *80*, 331–337.

Lara Pizarro, A.V. and E.Y. Park (2003). Lipase-catalyzed production of biodiesel fuel from vegetable oils contained in waste activated bleaching earth. Process Biochem. *38*, 1077–1082.

Lara, P.V. and E.Y. Park (2003). Lipase-catalyzed production of biodiesel fuel from vegetable oils contained in waste activated bleaching earth. Process Biochem. *38*, 1077–1082.

Lara, P.V. and E.Y. Park (2004). Potential application of waste activated bleaching earth on the production of fatty acid alkyl esters using *Candida cylindracea* lipase in organic solvent system. Enzyme Microb. Technol. *34*, 270–277.

Lee, K.-T., T.A. Foglia and K.-S. Chang (2002). Production of alkyl ester as biodiesel from fractionated lard and restaurant grease. J. Am. Oil Chem. Soc. *79*, 191–195.

Mittlebach, M. (1990). Lipase catalyzed alcoholysis of sunflower oil. J. Am. Oil Chem. Soc. *67*, 168–170.

Nelson, L.A., T.A. Foglia and W.N. Marmer (1996). Lipase-catalyzed production of biodiesel. J. Am. Oil Chem. Soc. *73*, 1191–1195.

Noureddini, H., X. Gao and R.S. Philkana (2005). Immobilized *Pseudomonas cepacia* lipase for biodiesel fuel production from soybean oil. Biores. Technol. *96*, 769–777.

Pandey, A., S. Benjamin, C.R. Soccol, P. Nigam, N. Krieger and V. T. Soccol (1999). The realm of microbial lipases in biotechnology. Biotechnol. Appl. Biochem. *29*, 119–131.

Pedersen, S. and M.W. Christensen (2000). Immobilized biocatalysts. Applied biocatalysis. P. Adlercreutz. Amsterdam, Harwood Academic Publishers: 213–228.

Reetz, M.T. (1997). Entrapment of biocatalysts in hydrophobic sol-gel materials for use in organic chemistry. Adv. Mater. *9*, 943–953.

Salis, A., M. Monduzzi and V. Solinas (2005a). Enzymes for biocatalysis in non aqueous media. Biocatalysis: Chemistry and Biology. A. Trincone. Kerala, Research Signpost. 29–53.

Salis, A., M. Pinna, M. Monduzzi and V. Solinas (2005b). Biodiesel production from triolein and short chain alcohols through biocatalysis. J. Biotechnol. *119*, 291–299.

Salis, A., E. Sanjust, V. Solinas and M. Monduzzi (2003a). Characterisation of Accurel MP1004 polypropylene powder and its use as a support for lipase immobilisation. J. Mol. Catal. B: Enzym. *24–25*, 75–82.

Salis, A., E. Sanjust, V. Solinas and M. Monduzzi (2005c). Commercial lipase immobilization on Accurel MP1004 porous polypropylene. Biocatal. Biotransf. *23*, 381–386.

Salis, A., I. Svensson, M. Monduzzi, V. Solinas and P. Adlercreutz (2003b). The atypical lipase B from *Candida antarctica* is better adapted for organic media than the typical lipase from *Thermomyces lanuginosa*. Biochim. Biophys. Acta *1646*, 145–151.

Samukawa, T., M. Kaieda, T. Matsumoto, K. Ban, A. Kondo, Y. Shimada, H. Noda and H. Fukuda (2000). Pretreatment of immobilized *Candida antarctica* lipase for biodiesel fuel production from plant oil. J. Biosci. Bioeng. *90*, 180–183.

Selmi, B. and D. Thomas (1998). Immobilized lipase-catalyzed ethanolysis of sunflower oil in a solvent free medium. J. Am. Oil Chem. Soc. *75*, 691–695.

Shah, S., S. Sharma and M.N. Gupta (2003). Enzymatic transesterification for biodiesel production. Indian J. Biochem. Biophys. *40*, 392–399.

Shah, S., S. Sharma and M.N. Gupta (2004). Biodiesel preparation by lipase-catalyzed transesterification of jatropha oil. Energy and Fuels *18*, 154–159.

Shieh, C.-J., H.-F. Liao and C.-C. Lee (2003). Optimisation of lipase-catalyzed biodiesel by response surface methodology. Biores. Technol. *88*, 103–106.

Shimada, Y., Y. Watanabe, T. Samukawa, A. Sugihara, H. Noda, H. Fukuda and Y. Tominaga (1999). Conversion of vegetable oil to biodiesel using immobilized *Candida antarctica* lipase. J. Am. Oil Chem. Soc. *76*, 789–793.

Shimada, Y., Y. Watanabe, A. Sugihara and Y. Tominaga (2002). Enzymatic alcoholysis for biodiesel fuel production and application of the reaction to oil processing. J. Mol. Catal. B: Enzym. *17*, 133–142.

Soumanou, M.M. and U.T. Bornscheuer (2003a). Improvement in lipase-catalyzed synthesis of fatty acid methyl esters from sunflower oil. Enzyme Microb. Technol. *33*, 97–103.

Soumanou, M.M. and U.T. Bornscheuer (2003b). Lipase-catalyzed alcoholysis of vegetable oils. Eur. J. Lipid Sci. Technol. *105*, 656–660.

Tuter, M., H.A. Aksoy, E.E. Gilbaz and E. Kursun (2004). Synthesis of fatty acid esters from acid oils using lipase B from *Candida antarctica*. Eur. J. Lipid Sci. Technol. *106*, 513–517.

Vermue, M.H. and J. Tramper (1995). Biocatalysis in non-conventional media: medium engineering aspects. Pure Appl. Chem. *67*, 345–373.

Watanabe, Y., Y. Shimada, A. Sugihara, H. Noda, H. Fukuda and Y. Tominaga (2000). Continuous production of biodiesel fuel from vegetable oil using immobilized *Candida antarctica* lipase. J. Am. Oil Chem. Soc. *77*, 355–360.

Watanabe, Y., Y. Shimada, A. Sugihara and Y. Tominaga (2001). Enzymatic conversion of waste edible oil to biodiesel fuel in a fixed-bed bioreactor. J. Am. Oil Chem. Soc. *78*, 703–707.

Watanabe, Y., Y. Shimada, A. Sugihara and Y. Tominaga (2002). Conversion of degummed soybean oil to biodiesel fuel with immobilized *Candida antarctica* lipase. J. Mol. Catal. B, Enzym. *17*, 151–155.

Wu, W.H., T.A. Foglia, W.N. Marmer and J.G. Phillips (1999). Optimizing production of ethyl esters of grease using 95% ethanol by response surface methodology. J. Am. Oil Chem. Soc. *76*, 517–521.

Xu, Y., W. Du and D. Liu (2005). Study on the kinetics of enzymatic intersterification of triglycerides for biodiesel production with methyl acetate as the acyl acceptor. J. Mol. Catal. B: Enzym. *32*, 241–245.

Xu, Y., W. Du, D. Liu and J. Zeng (2003). A novel enzymatic route for biodiesel production from renewble oils in a solvent-free medium. Biotechnol. Lett. *25*, 1239–1241.

Xu, Y., W. Du, J. Zeng and D. Liu (2004). Conversion of soybean oil to biodiesel fuel using lipozyme TL IM in a solvent-free medium. Biocatal. Biotransf. *22*, 45–48.

Yesiloglu, Y. (2004). Immobilized lipase-catalyzed ethanolysis of sunflower oil. J. Am. Oil Chem. Soc. *81*, 157–160.

Zhang, Y., M.A. Dubé, D.D. McLean and M. Kates (2003). Biodiesel production from waste cooking oil: 1. Process design and technological assessment. Biores. Technol. *89*, 1–16.

CHAPTER 20

USE OF LIPASES IN THE SYNTHESIS OF STRUCTURED LIPIDS IN SUPERCRITICAL CARBON DIOXIDE

JOSÉ DA CRUZ FRANCISCO[1], SIMON P. GOUGH[2] AND ESTERA S. DEY[1*]

[1]*Department of Pure and Applied Biochemistry and*
[2]*Department of Biochemistry, Lund University, Lund, Sweden*
[*]*estera.dey@tbiochem.lth.se*

1. INTRODUCTION

Interesterification is one of the techniques used to obtain oils or fats with desirable physicochemical characteristics. Additionally, interesterification can produce oils/fats with useful nutritional and health attributes such as the structured fats. The interesterification can be performed both chemically and enzymaticaly. However, enzymatic interesterification of oils and fats using immobilized lipases has attracted especial attention because of the lower operation temperature required and the yield of a more defined product.

In recent years a more environmentally accepted reaction medium for such reactions has been sought. This focused attention on enzymatic interesterification in supercritical fluids, particularly supercritical carbon dioxide ($scCO_2$), to replace organic solvents. $scCO_2$ as a reaction medium principally offers the same advantages for lipase catalysis as organic solvents. Hydrophobic lipid substrates are soluble, immobilized lipase is easily recovered and the reversal of hydrolysis favours the synthesis (Martins *et al*., 1994; Gunnlaugsdottir and Sivik, 1997). However $scCO_2$ provides even more advantages, see section 6.1.

This chapter discusses the use of lipase in $scCO_2$ for the transesterification reactions and explores the advantages, drawbacks and difficulties encountered in the attempt to produce structure oils/fats.

2. STRUCTURED LIPIDS

In this chapter, the terms lipid or structured lipids will be restricted to triacylglycerols which include oils and/or fats commonly used for food purposes. Structured lipids (SL) are defined as triacylglycerols containing medium (MCFA) or short-chain fatty

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 341–354

acids (SCFA) and long-chain essential/functional fatty acids produced by chemical modification or enzymatic synthesis. Structured lipids thus provide both a swift energy source as well as essential fatty acids. The idea of using structured lipids for nutritional applications arose from successful clinical applications of medium-chain triacylglycerols. Further research showed that the inclusion of both medium chain fatty acids and essential fatty acids in the same triacylglycerol improved the nutritional effectiveness of SL (Xuebing *et al*., 2002). Before absorption, triacylglycerols containing long chain fatty acids (LCFA) (>12 carbon atoms) must be hydrolyzed in the intestinal lumen. Subsequently, they must be re-synthesized in the enterocytes and then used to form lipoprotein particles known as chylomicrons. From the chylomicrons, long chain fatty acids (LCFAs) are released by serum lipases and incorporated in the tissues, stored in the adipocytes or burnt in the mitochondria. However, before uptake across the mitochondria membrane, LCFA must be converted to acylcarnitine by the acyltransferase. In contrast triacylglycerols containing MCFA can directly be absorbed without hydrolysis and preferentially transported through the portal venous system to the liver. Thus, MCFA are readily burnt via mitochondrial β-oxidation whereas most LCFA are incorporated into triacylglycerols in the hepatocyte. So, dietary MCFA intervention bypasses lipid deposition into adipocytes, the lipase-regulated metabolic pathway of LCFA. This can reduce the tendency to obesity (Mu *et al*., 2005). When 22% n-3 long-chain polyunsaturated fatty acids (n-3 LCPUFA) are also included in the diet endogenous oxidation of LCFA has also been demonstrated (Beermann *et al.*, 2003). Compared to LCFA, fats containing some MCFA, have been shown to reduce body weight in rats (Shinohara *et al*., 2005). MCFAs are metabolized as rapidly as glucose in the body. Because they are not readily re-esterified into triacylglycerols, there is little tendency to deposit as stored fat. MCFAs are therefore useful in the control of obesity. However, they can potentially raise serum cholesterol levels. MCFAs in a structured lipid substituted to glycerol at position sn-1, and sn-3 are better and more useful. Such lipid construction combines their inherent mobility, solubility, and easy metabolism. With beneficial polyunsaturated fatty acids located in position sn-2 it is possible to achieve good nutritional properties (Osborn and Akoh, 2002).

The composition of the triacylglycerols in each fat or oil is specific to its origin. The physical properties of the various fats/oils are directly related to the structure and distribution of the fatty acids in the triacylglycerol backbone (Parodi, 1982; Macrae, 1985; Christie, 1986; Willis *et al*., 1998). Naturally occurring fats or oils are mixtures of triacylglycerols. In such mixtures triacylglycerols with an ideal distribution of the fatty acids in the glycerol backbone corresponding to a structured lipid are not common and therefore they must be obtained artificially.

3. SYNTHESIS OF STRUCTURED LIPIDS VIA LIPASE

Structured lipids can be produced by chemical or enzymatic interesterification. Whether chemically or enzymaticaly, during interesterification there is a redistribution of acyl groups of the triacylglycerols participating in the reaction to achieve the desired structures (Willis and Marangoni, 1998).

The first studies on interesterification were published early in the 18th century. At the present interesterification plays an important role in the production of low caloric fats replacers (Smith *et al*., 1994; Rousseau and Marangoni, 1998a). Interesterification include; reactions between a triacylglycerol and a fatty acid (acidolysis), between an alcohol and a triacylglycerol (alcoholysis), glycerolysis (when glycerol acts as the alcohol) and between two triacylglycerols (transesterification) (Feuge, 1962; Sonntag, 1982; Malcata *et al*., 1990; Willis and Marangoni, 1998).

Transesterification is the most common type of interesterification in food applications. Chemical interesterification has been used for many years but research has presently directed to replace it by enzymatic interesterification. There are two reasons for this. Chemical interesterification requires severe reaction conditions and produces positional randomization of the acyl groups in the triacylglycerols obtained. Conversely, many factors recommend enzymatic interesterification. Enzymes work well under mild conditions, specifically produce well defined products and produce less waste. Reusable immobilized enzymes lead to more economical processes (Goh *et al*., 1993).

4. MOLECULAR STRUCTURE AND MECHANISM OF LIPASES

Lipases can be divided into two general structural classes: those which exist in an inactive form, where the active site is covered by a lid and those where the active site is always exposed to solvent. Examples of the commercially available classes of such enzymes are presented in Table 1. The lipozyme of *Mucor miehei* RM-IM is one of the lipases with movable lid which have been used in synthesis purposes in supercritical carbon dioxide. This lipase from Novozymes, has been immobilized on a weak-anion-exchange resin (IM) (Boel *et al*., 1988). It can rearrange the fatty acids in the *sn*-1 and *sn*-3 positions of triglycerides. Only the *sn*-2 positions are preserved. It belongs to the alpha/beta hydrolase fold enzyme family with an active site triad of serine aspartate and histidine similar to serine proteases (Holmquist, 2000). Kinetic experiments suggested that lipase activation involved the formation of a local chamber on phase contact area where the hydrolysis occurs. Three regions were suggested in the active centre of lipolytic enzymes: 1) a region responsible for the recognition" of substrate phase surface; 2) a binding region, participating in hydrophobic interaction with a single substrate molecule, located in the insoluble phase; and 3) catalytical region (Chapus and Semeriva, 1976).

The crystal structure of *Mucor miehei* lipase (Brzozowski *et al*., 1991) shows that the serine esterase inhibitor *p*-nitrophenyl diethyl phosphate has reacted with the S144 to give a serine bound diethyl phosphate. This identifies S144 as the active site serine. The C-terminal part of the un-liganded lipase structure can be superimposed on the liganded structure (Fig. 1). However the N-terminal short amphiphilic helix (residues 85 to 91) acts as a "closed hydrophobic lid" in the unliganded structure. The hydrophobic residues (white) of the amphiphilic helix are internalised and the polar residues (black) solvent exposed (Fig. 1 top left). The hydrophobic residues

Table 1. Commercially available lipases used in synthesis of structure lipids

Commercial name	Lipase	Lipase class
Novozym®435	*Immobilized Candida antarctica lipase B*	Permanently open active site
Lipozyme RM IM®	*Mucor miehei immobilized*	Active site covered by movable lid
Lipozyme® TM IM	*Thermomyces lanuginosus* on porous silica	Active site covered by movable lid
Lipex®100L	*Thermomyces lanuginosus* Mutant with better lipid susface adsorption (Snab, 2004)	Active site covered by movable lid Stable to detergents
Palatase®20000 L,	*Mucor miehei*(Barron *et al*, 2004)	Active site covered by movable lid
Novozym CALB L®	*Candida antarctica* lipase B	Permanently Open active site
Lipozyme® TL100 L	*Thermomyces lanuginosus*	Active site covered by movable lid

lie above the active site histidine and serine. This lid is inverted in the inhibitor bound structure. The helix extended by the hydrophobic residue 92, moves away from the catalytic triad and turns outward to face the solvent to give the "open" structure (Fig. 1 top right). The active site residues can be viewed from above as in Figs. 2 and 3, surrounded by the solvent accessible (>30%) residues. It can be clearly seen that non-polar residues are clustered on the top compared to the bottom of these pictures. In the orientation of these pictures the active site is partially buried on top by the internalised (<30% accessible) hydrophobic residues I204 and V205 whereas on the bottom, solvent access to active site histidine and aspartate is unrestricted (not shown). The main difference in the liganded structure (Fig. 1 top right), compared to the unliganded (Fig. 1 top left), is the presence of the non-polar solvent-exposed patch formed by hydrophobic residues in the amphiphilic helix. This patch is considered to facilitate contact with lipid droplets allowing product release and substrate binding. In the liganded structure the strongly conserved W88 of the amphilic helix, although exposed to solvent, remains close to the hydrophobic ethyl groups of the inhibitor and presumably interacts hydrophobically with the non-polar. A completely analogous behaviour of the corresponding amphiphilic helix is seen in the structure of the active site S144A mutant *Thermomyces lanuginosa* lipase. In this case the open lid structure is seen with bound oleic acid and the closed lid with the unliganded enzyme (Brzozowski *et al*, 2000). These fungal lipases are assigned to the class Lipase 3. Conservation of the catalytic triad and sequence similarity of lipases from these fungi and from organisms including viruses, bacteria, plants and other eukaryotes including plasmodium and homo suggests that these all have a similar structure and mechanism.

A simple two state model of the lipase suggested that closed state might be more stable in aqueous media rendering the active site inaccessible to water soluble substrates. However the lid of the *Humicola lanuginosa* (now called *Thermomices*

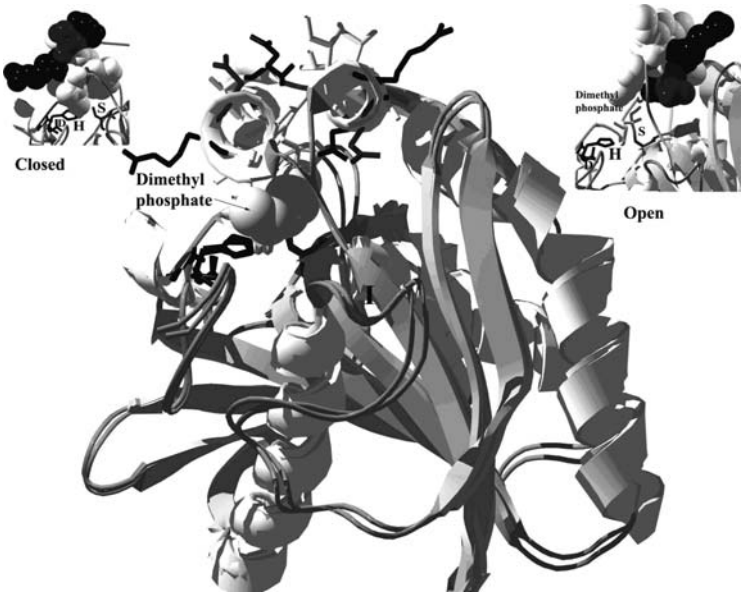


Figure 1. Superinposed of unliganded and liganded (bound to *p*-nitrophenyl diethyl phosphate) structures of *Mucor miehei* lipase. Top left: the hydrophobic residues (white) of the amphiphilic helix are internalised and the polar residues (black) solvent exposed. Top right: the helix moves away from the catalytic triad and turns outward to face the solvent

langinosus) lipase is disordered irrespective of the media's ionic strength. *Rhizopus delemar* lipase crystallized in detergent exists in an asymmetrica dimeric form and both open and closed lid forms are seen. These new results call into question the

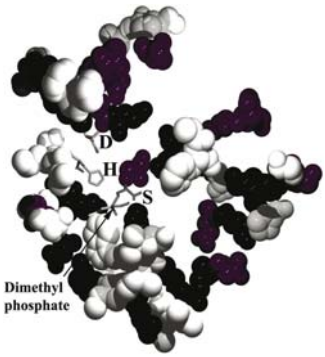


Figure 2. Liganded (bound to *p*-nitrophenyl diethyl phosphate) open structure of *Mucor miehei* lipase. Hydrophobic and hydrophilic residues are represented in white and black, respectively

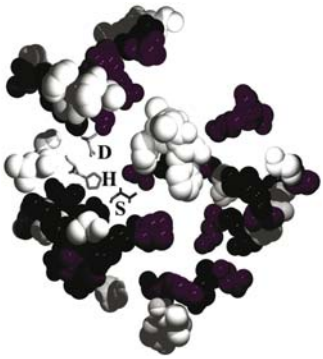


Figure 3. Unliganded closed structure of *Mucor miehei* lipase. Hydrophobic and hydrophilic residues are represented in white and black, respectively

simplicity of the enzyme theory of interfacial activation of *Penicillium camemberti* and *Rhizopus oryzae* lipases (Derewenda *et al*., 1994a, b).

The structure of a mutant M211S/R215L with changed substrate specificity shows an unexpected dimerization and domain sapping involving the N-terminal region. When a tailored substrate-like irreversible inhibitor 1-hexadecanesulfonyl chloride binds, a cis-trans isomerization of the F37-P38 peptide bond negates this dimerization. Different structural rearrangements of an open to a closed form are a general property of other lipases (DeSimone *et al*., 2004a). Kinetic analysis suggests that short chain p-nitrophenyl-hexanoate substrates are metabolised by a different kinetic mode than long chain ones e. g. p-nitrophenyl dodecanoate. This was confirmed by different binding modes in the structures(DeSimone *et al*., 2004b). The main conclusion is that, although it would be expected that the active form of the movable lid lipase be more stable in solvent of lower ionic strength, organic solvents or $scCO_2$, this remains to be conclusively demonstrated.

5. REACTIONS INVOLVING LIPASE IN $ScCO_2$

Since Randolph *et al*. (Randolph *et al*., 1985) and Nakamura *et al*., (Nakamura *et al*., 1985) established the stability and activity of enzymes in supercritical fluids, several lipase reactions in $scCO_2$ have been described. Examples include esterification of free fatty acids and primary alcohols (Steytler *et al*., 1991; Marty *et al*., 1992b; Yu *et al.*, 1992), alcoholysis of propylacetate with geraniol (Chulalaksananukul *et al*., 1993) and cod liver oil and ethanol (Gunnlaugsdottir and Sivik, 1995) as well as the acidolysis of triolein and stearic acid using a 1,3-specific regiospecific lipase (Chi *et al.*, 1988; Nakamura, 1994). Transesterification of triolein with ethylbehemate by immobilized lipase (Yoon *et al*., 1996) as well as the transesterification of different triglycerides (milk fat and canola oil) (Yu *et al*., 1992) were also reported.

6. FACTORS INFLUENCING LIPASE ACTIVITY AND SYNTHESIS IN $ScCO_2$

The pH, temperature, water content and activity, enzyme purity and presence of other proteins, substrate composition and steric hindrance, surface active agents and product accumulation are factors that can influence the lipase activity during interesterification. These factors have been exhaustively discussed in an elegant review by Willis and Marangoni (Willis and Marangoni, 1998). However, this review does not include processes where supercritical fluid is the reaction medium. Many studies involving interesterification in $scCO_2$ have been performed and the factors affecting the reaction in organic media have been investigated.

A fluid is considered to be supercritical when its temperature and pressure are above the critical point; the maxima pressure and temperature combination in which the equilibrium gas-liquid is possible. Because supercritical fluids are good solvents, they have received great attention as possible new solvents for enzymatic reactions since the mid-1980's (Nakamura, 1990; Aaltonen, 1991; Marty *et al.*, 1992a). This is because supercritical fluids have physical chemical properties between those of liquid and gas. The high diffusivities, lower viscosities and lower surface tension relative to liquids, accelerate the reaction rates. Low viscosity and high diffusivity of the reaction medium, for instance, speed up reactions controlled by mass transfer (Stahl *et al*, 1988); *e.g.* lipase in $scCO_2$ is about 8–9 fold more efficient than in hexane has been reported. In $scCO_2$ Mucor *miehei* immobilized lipase esterified mystiric acid and ethanol with an average efficiency of 38–68% compared to 4–9% in hexane. This large increase in efficiency was attributed to the superior rate of diffusion of mysteric acid in $scCO_2$ where its diffusion coefficient is hundred times higher than in hexane (Bernard *et al*, 1995). Because of its relatively low critical point (31 °C; 7.3MPa), $scCO_2$ has been widely explored as reaction medium. A low critical temperature allows enzymatic reactions to occur even at physiological temperatures; a good feature for biocatalysis. Furthermore, the readily available CO_2 is non-volatile, non-flammable, and non-toxic. $scCO_2$ has excellent properties for lipase reactions (Nakamura, 1990; Aaltonen and Rantakyla, 1991; Russell *et al*, 1991) and the possibility of changing the physical state by simple manipulation of the pressure and temperature facilitates product recovery. This makes $scCO_2$ specially attractive for food and pharmaceutical industrial processes (Marty *et al*, 1994) and it is environmentally accepted (Francisco *et al*, 2003a).

6.1. Enzyme Stability in $ScCO_2$

Enzymes are proteins. Proteins are practically insoluble in supercritical CO_2. However, enzyme stability must be considered since in esterification and interesterification reactions, water, heat and additionally when using supercritical CO_2, pressure are involved.

In $scCO_2$ pressure is not a threat to protein configuration since it is only irreversibly influenced at pressure much higher (600 MPa and higher) than those

used (Bridgman, 1914; Aoki *et al*., 1968; Murphy, 1978; Chryssomallis *et al*., 1981; Kornblatt *et al*., 1982; Kornblatt *et al*., 1986). When protein denaturation was observed at lower pressures and temperatures it has been attributed to water content and heat rather than pressure (Weder, 1990).

After six days of incubation at pressure from 13 to 18MPa at 40 °C, Marty and co-workers (Marty *et al*., 1990) observed only a 10% decay of the enzyme activity. At 9.6MPa and 70 °C, Lipozyme RM IM retained complete activity for more than 26 hours (Bisten, 2004). Loss of enzyme activity during the depressurization step has been reported (Kasche *et al*., 1988; Marty *et al*., 1992b), however this effect was only observed in hydrated $scCO_2$.

The good stability of the enzyme makes this technique attractive in the pharmaceutical field where mild operation temperatures and solvent safety is crucial.

6.2. Water Activity

Water influences the hydration of the enzyme and consequently the conformation and optimal activity (Rupley *et al*., 1983; Zaks *et al*., 1988; Klibanov, 1989; Chulalaksananukul *et al.*, 1990; Hirata *et al*., 1990; Chulalaksananukul *et al*., 1992; Halling, 1994). The water content of the solid support is deemed to be the most important aspect for the enzyme activity. In esterification, water is one of reaction products whose effect in the thermodynamic of the reaction is predominant (Marty *et al*., 1992a).

Chulalaksananukul and co-workers (Chulalaksananukul *et al*., 1993) have investigated the effect of addition of water and increased temperature (40 to 100 °C at 14MPa) on the stability of immobilized lipase from *M. miehei* in $scCO_2$. The more water added and the higher the temperature applied, the lower the residual activity of the enzyme. For instance, by adding 30μl to 100mg of enzyme, the residual activity dropped from 70% at 40 °C to 30% at 100 °C; *c. f*. Fig. 2 (Chulalaksananukul *et al.*, 1993). The decrease of enzyme activity with the water increase was evident for every temperature tested. From the same results it was observed that even without water addition, the enzyme activity decreased solely as a consequence of temperature increase. The authors also investigated the influence of the added water on the lipase catalyzed interesterification of geraniol and propyl acetate. The initial rate increased with increasing the amounts of water added to a maximum value at 3μl and 10μl of water in hexane and $scCO_2$, respectively and then a progressive decreasing was observed. The difference between the optimum water amounts in both solvents was attributed to the fact that $scCO_2$ is less hydrophobic than hexane. Quantification and comparison of the effect of water content in organic solvent and $scCO_2$ and its relationship to pressure and temperature have been investigated (Marty *et al*., 1992b).

6.3. Temperature and Pressure

The physical properties of $scCO_2$ depend greatly on the pressure and temperature. It is conveniently possible to tune its solvent power and thus its behavior as a

reaction medium. Yoon and co-workers (Yoon *et al*., 1996) studied the effect of pressure and temperature on the transesterification of triolein and ethylbehenate by immobilized lipase Lipozyme IM. This enzyme is thermo stable in $scCO_2$ below 80 °C (Marty *et al*., 1990; Chulalaksananukul *et al*., 1993) therefore the study could be performed in the range of 40–70 °C. The hydrolytic and transesterification activity increased with increasing temperature. The increase in temperature enables reactants to cross the activation energy barrier of the transition state (Whitaker, 1994). In $scCO_2$ the increase in temperature increases the solubility of triolein and ethylbehenate and decreases the density. The lower density and viscosity of the $scCO_2$ results in an increase in the mass transfer rate of the substrates and products in the immobilized enzyme. The overall conversion reached its maximum between 50 and 60 °C.

The effect of pressure was also analyzed for the same system at 50 °C. The maximum transesterification activity was observed at 15 MPa while the hydrolytic activity reached its maximum at 12 and 25 MPa. At lower pressures the rates of both transesterification and hydrolysis decreased due to the low substrate solubility at the temperature studied. The solubility of the substrates however increase with increasing pressure at constant temperature (Chrastil, 1982) which may result in the increase in the reaction rate. However, pressure increase can result in decreased rates as reported for other systems (Erickson *et al*., 1990; Saito *et al*., 1994).

6.4. Chain Length

The chain-length substrate selectivity of various lipases has been studied in a number of works (Janssen *et al*., 1996; Arsan *et al*., 2000). Among these, the report by Vaysse and co-corkers (Vaysse *et al*., 2002) represents an exhaustive description of the selectivity of various enzymes to different chain length substrates for hydrolysis, esterification and alcoholysis in aqueous medium. $scCo_2$ was not the reaction medium but it was shown that the transesterification of the triglyceride with the ester of the fatty acid may be more effective than that with the free fatty acid. This was attributed to the greatly decreased solubility of the free fatty acid in with its increase in molecular weight (Yoon *et al*., 1996).

6.5. Continuous Synthesis

The continuous fixed bed reactor has been reported as being the most suitable for development of new industrial applications *e.g.* ester synthesis by reversion of hydrolysis (Balcão *et al*., 1996). The Lipozyme catalyzed esterification of oleic acid with ethanol or the transesterification of triolein with ethanol in n-hexane in a continuous packed-bed reactor showed a drastic loss of enzymatic activity with time (Marty *et al*., 1997). This loss was attributed to the adsorption of polar substances, water, or glycerol onto the enzyme support. Solutions advanced include the creation of a hydrophilic layer that could prevent the diffusion of the hydrophobic substances from the medium to the enzyme or the modification of the optimal thermodynamic

water on the enzyme. Colombié and co-workers (Colombie *et al*., 1998) have proposed controlling thermodynamic water activity in the fixed bed reactor during continuous esterification of oleic acid with ethanol in order to maintain a high and constant conversion. In continuous esterification in $scCO_2$ the water released as a reaction product is continuously removed by the $scCO_2$ stream keeping the necessary water activity for the enzyme almost unchanged over the reaction process. While the water removal drives the esterification to completion, for transesterification this is not so. In transesterification reactions water is produced. Water has some solubility in $scCO_2$ (Wiebe *et al*., 1941; King *et al*., 1983; Francisco *et al*., 2003b) therefore in continuous transesterification the enzyme is dried by the $scCO_2$ and consequently inactivated. This problem has been solved by saturating glyceryltrioleate with water prior to the interesterification with caprylic acid (Bisten, 2004). The addition of water during esterification reactions has been earlier reported earlier for batch and continuous trials in both n-hexane and $scCO_2$ (Dumont *et al*., 1991; Dumont *et al*., 1992; Bernard and Barth, 1995).

The effect of flow rate on the product quality of a reaction was illustrated by Gunnlaugsdottir and Sivik (Gunnlaugsdottir and Sivik, 1997) for the continuous lipase catalyzed transesterification of cod liver oil with ethanol. In a continuous reaction by selectively extracting the desired product or by-product, the higher the flow rate, the higher the recovery of the ethyl esters. Such behavior is due to the fact that the reaction mixture is exposed to a higher volume of extraction fluid for a set period of time.

The continuous transesterification of natural occurring vegetable oils to obtain structured oils faces other challenges. Because natural occurring oils are mixtures of different triacylglycerols the range of products yielded is wide and the overall yield of the desired compounds is reduced. On the other hand, the purification of the reaction product is rendered complex. One possibility to overcome such constrain is to develop a reaction/fractionation process in a system with a series of cyclone separator downstream or perform the fractionation in a fixed bed column with reflux. The use of a series of cyclone separator system was advanced by Knez and collaborators (Knez and Habulin, 1992 and 1994; Knez *et al*., 1995). This represents one of the present limitations and challenges in the use of $scCO_2$ for producing structured lipids and the challenge.

7. $ScCO_2$ VERSUS ORGANIC MEDIUM

The choice for the $scCO_2$ as the reaction medium will still be a matter of debate. $scCO_2$ offer the same advantages for lipase catalysis as organic solvents namely the solubilization of the hydrophobic lipid substrates, simple recovery of immobilized lipase as well as the reversal of hydrolysis in favor of synthesis (Martins *et al*., 1994; Gunnlaugsdottir and Sivik, 1997). $scCO_2$ will also support reactions thermodynamically unfavorable in aqueous media (*e.g.* synthesis of esters and amides) (Martins *et al.*, 1994). Besides the advantageous $scCO_2$ properties pointed out earlier in this chapter, this solvent allows combined processes reaction/extraction and fractionation

of the products. Some drawbacks should be noted. For instance, the synthesis of geranyl acetate was performed in $scCO_2$ by immobilized *M. miehei* lipase catalyzed transesterification of geraniol and propyl acetate at 14MPa and 40 °C. The yield of geranyl acetate was found to be 30% after three days (Chulalaksananukul *et al*, 1993). In a previous work however, the yield of the same reaction in hexane was 85% (Chulalaksananukul *et al*, 1992) after the same period of time. The great difference in the reaction yield in both media is attributed to reactants and product solubility effect (Halling, 1990). When the solvent is less polar the solvation of the more polar OH and COOH group decline relatively faster than that of the ester group. $scCO_2$ is more polar than n-hexane, thus in case of esterification the equilibrium is favored in n-hexane where the product is less soluble.

In general the use and exploration of $scCO_2$ as a reaction medium will have to be considered for each case of interest. More effort is needed until the time when commercial plants can be constructed for structured lipids production, especially in continuous processes in $scCO_2$.

REFERENCES

Aaltonen, O. and M. Rantakyla (1991). Biocatalysis in supercritical CO_2. Trends Biotechnol. *21*(4), 240–248.

Aoki, K., K. Hiramatsu, M. Tanaka and S. Kaneshina (1968). Bovine serum albumin exposed to high pressure. Biochim. Biophys. Acta *160*, 168–177.

Arsan, J. and K.L. Parkin (2000). Selectivity of *Rhizomucor miehei* lipase as affected by choice of cosubstrate system in ester modification reactions in organic media. Biotech. Bioeng *69*, 221–226.

Balcão, V.M., A.L. Paiva and F.X. Malcata (1996). Bioreactors with immobilized lipases: State of the art. Enzyme Microb. Technol. *18*, 392–416.

Barron, L., I. Hernandez, A. Bilbao, C.E. Flanagan, A.I. Najera, M. Virto, F.J. Perez-Elortondo, M. Albisu and M. de Renobales (2004). Changes in lipid fractions and sensory properties of Idiazabal cheese induced by lipase addition. Journal of Dairy Research *71*, 372–379.

Beermann, C., J. Jelinek, T. Reinecker, A. Hauenschild, G. Boehm and H.-U. Klör (2003). Short term effects of dietary medium-chain fatty acids and n-3 long-chain polyunsaturated fatty acids on the fat metabolism of healthy volunteers. Lipids in Health and Desease *2*, 1–10.

Bernard, P. and D. Barth (1995). Internal mass transfer limitation during enzymatic esterification in supercritical carbon dioxide and hexane. Biocatalysis and Biotransform *12*, 299–308.

Bister, T. (2004). Transesterification in supercritical carbon dioxide. (Master Thesis). Department of Pure and Applied Biochemistry. Lund, Lund University, *35*.

Boel, E., B. Huge-Jensen, M. Christensen, L. Thim and N.P. Fiil (1988). Rhizopus Miehei triglyceride lipase is synthensized as a percursor. Lipids *23*, 701–706.

Bridgman, P.W. (1914). The coagulation of albumin by pressure. J. Biol. Chem. *19*, 511–512.

Brzozowski, A.M., U. Derewenda, Z.S. Derewenda, G.G. Dodson, D. M. Lawson, J.P. Turkenburg, F. Bjorkling, B. Huge-Jensen, S.A. PAtkar and L. Thim (1991). A model for interfacial activation in lipases from the structure of fungal lipase-inihibitior complex. Nature *351*, 491–494.

Brzozowski, A.M., H. Savage, C.S. Verma, J.P. Turkenburg, D.M. Lawson, A. Svendsen and S. Patkar (2000). Structural origins of the interfacial activation in *Thermomyces (Humicola) lanuginosa* lipase. Biochem. *39*, 15071–15082.

Chapus, C. and M. Semeriva (1976). Mechanism of pancreatic lipase action. 2. Catalytic properties of modified lipases. Biochem. *15*, 4988–4991.

Chi, Y.M., K. Nakamura and T. Yano (1988). Enzymatic interesterification in supercritical carbon dioxide. Agric. Biol. Chem. *52*(6), 1541–1550.

Chrastil, J. (1982). Solubility of solids and liquids in supercritical gases. J. Phys. Chem. *86*, 3016–3021.

Christie, W.W. (1986). The positional distributions of fatty acids in triglycerides. The Analysis of Oils and Fats. J.J. Hamilton and J.B. Rossel. London, Elsevier Appl. Sci., 313–339.

Chryssomallis, G.S., P.M. Torgerson, H.G. Drickamer and G. Weber (1981). Effect of hydrostatic pressure on lysozyme and chymotrypsinogen detected by fluorescence polarization. Biochem. *20*, 3955–3959.

Chulalaksananukul, W., J.S. Condoret and D. Combes (1992). Kinetics of geranyl acetate synthesis by lipase-catalyzed transesterification in n-hexane. Enzyme Microb. Technol. *14*, 293–298.

Chulalaksananukul, W., J.S. Condoret, P. Delorme and R.M. Willemot (1990). Kinetic study of esterification by immobilized lipase in n-hexane. FEBS Lett. *276*, 181–184.

Chulalaksananukul, W., J.-S. Conduret and D. Combes (1993). Geranyl acetate synthesis by lipase-catalyzed transesterification in supercritical carbon dioxide. Enzyme Microb. Technol. *15*, 691–698.

Colombie, S., R.J. Tweddell, J.-S. Condoret and A. Marty (1998). Water activity control: a way to improve the efficiency of continuous lipase esterification. Biotech. Bioeng *60*, 362–368.

Derewenda, U., L. Swenson, R. Green, Y. Wei, S. Yamaguchi, R. Joerger, M.J. Haas and Z.S. Derewenda (1994a). Current progress in crystallographic studies of new lipases from filamentous fungi. Protein. Eng. *7*, 551–557.

Derewenda, U., L. Swenson, Y. Wei, R. Green, P.M. Kobos, R. Joerger, M.J. Haas and Z.S. Derewenda (1994b). Conformational lability of lipases observed in the absence of an oil-water interface: crystallographic studies of enzymes from the fungi Humicola lanuginosa and Rhizopus delemar. J. Lipid Res. *35*, 524–534.

DeSimone, G., L. Mandrich, V. Menchise, V. Giordano, F. Febbraio, M. Rossi, C. Pedone and G. Manco (2004a). A Substrate-induced Switch in the Reaction Mechanism of a Thermophilic Esterase: Kinetic evidences and structural basis. J. Biol. Chem. *279*, 6815–6823.

DeSimone, G., V. Menchise, V. Alterio, L. Mandrich, M. Rossi, G. Manco and C. Pedone (2004b). The Crystal Structure of an EST2 Mutant Unveils Structural Insights on the H Group of the Carboxylesterase/Lipase Family. J. Biol. Chem. *343*, 137–146.

Dumont, T., D. Barth, C. Corbier, G. Branlant and M. Perrut (1992). Enzymatic reaction kinetic: comparison in organic solvent and in supercritical carbon dioxide. Biotech. Bioeng *39*, 329–333.

Dumont, T., D. Barth and M. Perrut (1991). Continuous synthesis of ethyl misterate by enzymatic reaction in supercritical carbon dioxide. Proceeding of the 2nd International Symposium on Supercritical Fluids, Boston, USA.

Erickson, J.C., P. Schyns and C.L. Cooney (1990). Effect of pressure on an enzymic reaction in a supercritical fluid. AIChE J. *36*, 299–301.

Feuge, R.O. (1962). Lectures of the 1962 Short Course on Developments in Fat Chemistry; the Am. Oil Chemists' Soc. J.W. Cowan (chairman). III. Derivatives of fats for use as foods. J. Am. Oil Chem. Soc. *39*, 521–527.

Francisco, J. d. C. and E.S. Dey (2003a). Supercritical fluids as alternative, safe, food-processing media: an overview. Acta Microb. Polonica *52*(Supl.): 35–43.

Francisco, J. d. C., C. Turner, D. Topgård, B. Sivik and B. Bergenståhl (2003b). Liquid Crystalline Properties and Extractability of Monoolein-Water Systems by Supercritical Carbon Dioxide. Colloids and Surfaces A *213*, 69–78.

Goh, S.H., S.K. Yeong and C.W. Wang (1993). Transesterification of cocoa butter by fungal lipases: effect of solvent on 1,3-specificity. J. Am. Oil Chem. Soc. *70*, 567–570.

Gunnlaugsdottir, H. and B. Sivik (1995). Lipase-catalyzed alcoholysis of cod liver oil in supercritical carbon dioxide. J. Am. Oil Chem. Soc. *72*, 399–405.

Gunnlaugsdottir, H. and B. Sivik (1997). Lipase-catalyzed alcoholysis with supercritical carbon dioxide extraction 1: Influence of flow rate. J. Am. Oil Chem. Soc. *74*, 1483–1490.

Halling, P.J. (1990). Solvent selection for biocatalysis in mainly organic systems: predictions of effects on equilibrium position. Biotech. Bioeng *35*, 691–701.

Halling, P.J. (1994). Thermodynamic predictions for biocatalysis in nonconventional media: theory, tests, and recommendations for experimental design and analysis. Enzyme Microb. Technol. *43*, 4.

Hirata, H., K. Higuchi and T. Yamashina (1990). Lipase-catalyzed transesterification in organic solvent: effects of water and solvent, thermal stability and some applications. J. Biotechnol. *14*, 157–167.

Holmquist, M. (2000). Alpha/beta-hydrolase fold enzymes: structures, functions and mechanisms. Current Protein Peptide Sci. *1*, 209–235.

Janssen, A.E.M., A.M. Vaidya and P.J. Halling (1996). Substrate specificity and kinetics of *candica rugosa* lipase in organic media. Enzyme Microb. Technol. *18*, 340–346.

Kasche, V., R. Schlothauer and G. Brunner (1988). Enzyme denaturation in supercritical CO_2: Stabilizing effect of S-S bonds during the depressurization step. Biotechnol. Lett. *10*, 569–574.

King, M.B., D.A. Alderson, F.H. Fallaha, D.M. Kassim, K.M. Kassim, J.R. Sheldon and R.S. Mahmud (1983). Some vapor/liquid and vapor/solid equilibrium measurements of relevance for supercritical extraction operations, and their correlation. Chemical Engineering at Supercritical Fluid Conditions. M.E. Paulaitis, J.M.L. Penninger, R.D. Gray, Jr. and P. Davidson. Ann Arbor, MI, Ann Arbor Science: 31–80.

Klibanov, A.M. (1989). Enzymic catalysis in anhydrous organic solvents. Trends Biochem. Sci. *14*, 141–144.

Knez, Z. and M. Habulin (1992). Lipase catalyzed esterification in supercritical carbon dioxide. Biocatalysis in Non-Concentional Media. J. Tramper, Elsevier Science Publishers: 401–407.

Knez, Z. and M. Habulin (1994). Lipase catalysed esterification at high pressure. Biocatalysis *9*, 115–121.

Knez, Z., V. Rizner, M. Habulin and D. Bauman (1995). Enzymatic synthesis of oleyl oleate in dense fluids. J. Am. Oil Chem. Soc. *72*, 1345–1349.

Kornblatt, J.A., A.M. English and G.H.B. Hoa (1986). The effects of pressure on yeast cytochrome c peroxidase. Eur. J. Biochem *159*, 39–43.

Kornblatt, J.A. and G.H.B. Hoa (1982). Conformations of cytochrome oxidase: thermodynamic evaluation of the interconversion of the 418- and 428-nm forms. Biochem. *21*, 5439–5444.

Macrae, A.R. (1985). Interesterification of fats and oils. Studies in Organic Chemistry (Amsterdam) *22*(Biocatal. Org. Synth): 195–208.

Malcata, F.X., H.R. Reyes, H.S. Garcia, C.G. Hill and C.H. Amundson (1990). Immobilized lipase reactors for modification of fats and oils - a review. J. Am. Oil Chem. Soc. *67*, 890–910.

Martins, J.F., I.B. d. Carvalho, T.C. d. Sampaio and S. Barreiros (1994). Lipase-catayzed enantioselective esterification of glycidol in supercritical carbon dioxide. Enzyme Microb. Technol. *16*, 785–790.

Marty, A., W. Chulalaksananukul, J.S. Condoret, R.M. Willemont and G. Durand (1990). Comparison of lipase-catalyzed esterification in supercritical carbon dioxide and n-hexane. Biotechnology Letters *12*, 11–16.

Marty, A., W. Chulalaksamanukul, J.-S. Condoret and D. Combes (1992a). Transesterification and esterification in supercritical carbon dioxide. Colloque INSERN-High pressure and Biotechnology. C. Balny, R. Hayashi, K. Heremans and P. Masson. *224*, 461–463.

Marty, A., W. Chulalaksananukul, R.M. Willemot and J.S. Condoret (1992b). Kinetics of lipase-catalyzed esterification in supercritical CO_2. Biotech. Bioeng *39*, 273–280.

Marty, A., D. Combes and J.-S. Condoret (1994). Continuous reaction-separation process for enzymatic esterification in supercritical carbon dioxide. Biotech. Bioeng *43*, 497–504.

Marty, A., V. Dossat and J.-S. Condoret (1997). Continuous operation of lipase-catalyzed reactions in nonaqueous solvents: influence of the production of hydrophilic compounds. Biotech. Bioeng *56*, 232–237.

Mu, H. and T. Porsgaard (2005). The metabolism of structured triacylglycerols. Prog. Lipid Res. *44*(6), 430–448.

Murphy, R.B. (1978). Anomalous stability of insulin at very high pressure. Experientia *34*, 188–189.

Nakamura, K. (1990). Biochemical reactions in supercritical fluids. Trends Biotechnol. *8*(10), 288–292.

Nakamura, K. (1994). Biochemical reactions in supercritical fluids. Supercritical Fluid Processing of Food and Biomaterials. S.S.H. Rizvi. Glasgow, UK, Blakie: 54–61.

Nakamura, K., Y.M. Chi, Y. Yamada and T. Yano (1985). Lipase activity and stability in supercritical carbon dioxide. Chem. Eng. Commun. *45*(1–6), 207–212.

Osborn, H.T. and C.C. Akoh (2002). Structured lipids-novel fats with medical, nutraceutical, and food applications. Comp. Rev. Food Sci. Food Safety *1*, 93–103.

Parodi, P.W. (1982). Positional distribution of fatty acids in the triglyceride classes of milk fat. J. Dairy Res. *49*(1), 73–80.

Randolph, T.W., H.W. Blanch, J.M. Praunitz and C.R. Wilke (1985). Enzymatic catalysis in a supercritical fluid. Biobetchnol. Lett. *7*, 325–328.

Rousseau, D. and A.G. Marangoni (1998). Chemical interesterification of food lipids: theory and practice. Food Sci. Technol. (N.Y.) *88*(Food Lipids: Chemistry, Nutrition and Biotechnology): 251–281.

Rupley, J.A., E. Gratton and G. Careri (1983). Water and globular proteins. Trends Biochem. Sci. *8*, 18–22.

Russell, A.J. and E.J. Beckman (1991). Enzymatic activity in supercritical fluids. Appl. Biochem. Biotechnol. *31*, 197–211.

Saito, N., O. Sato, Y. Ikushima, K. Hatakeda and S. Ito (1994). Productivity of isopropyllideneglycerol acyl ester synthesis by enzymatic reaction in supercritical carbon dioxide. Third Int. Symp. Supercrit. Fluids, Strasburg, France.

Shen, X.M., T.W. d. Loos and J. d. S. Arons (1992). Enzymatic reaction in organic solvents and supercritical gases. Biocatalysis in Non-Conventiional Media. J. Tramper, Elsevier Science Publishers: 417–423.

Shinohara, H., A. Ogawa, M. Kassai and T. Aoyama (2005). Effect of Randomly Interesterified Triacylglycerols Containing Medium- and Long-Chain Fatty Acids on Energy Expenditure and Hepatic Fatty Acid Metabolism in Rats. Biosci. Biotech. Biochem. *69*, 1811–1818.

Smith, R.E., J.W. Finley and G.A. Leveille (1994). Overview of SALATRIM: A family of low-calorie fats. J. Agric. Food Chem. *42*, 432–434.

Snabe, T. (2004). Mechanisms of Triacylglycerol Film Lipolysis. Biostructure and Protein engineering. Ålborg, Ålborg University: *202*.

Sonntag, N.O.V. (1982). Glycerolysis of fats and methyl esters - status, review and critique. J. Am. Oil Chem. Soc. *59*, 795A–802A.

Stahl, E., K.-W. Quirin and D. Gerard (1988). Dense Gases for Extraction and Refining. Berlin Heidlberg, Springer.

Steytler, D.C., P.S. Moulson and J. Reynolds (1991). Biotransformations in near-critical carbon dioxide. Enzyme Microb. Technol. *13*, 221–226.

Vaysse, L., A. Ly, G. Moulion and e. Dubreucq (2002). Chain-length selectivity of various lipases during hydrolysis, esterification and alcoholysis in biophasic aqueous medium. Enzyme Microb. Technol. *31*, 648–655.

Weder, J.K.P. (1990). Influence of supercritical carbon dioxide on proteins and amno acids - an overview. The Cafe, Cacao *34*, 87–96.

Whitaker, J.R. (1994). Principles of Enzymology for the Food Sciences. New York, Dekker (New York, N.Y.).

Wiebe, R. and V.L. Gaddy (1941). Vapor pressure composition of carbon dioxide-water mixtures at various temperatures and pressures to 700 atmospheres. J. Am. Chem. Soc. *63*, 475–477.

Willis, W.M. and A.G. Marangoni (1998). Enzymatic interesterification. Food Sci. Technol. (N.Y.) 88(Food Lipids, Chemistry, Nutrition and Biotechnology): 665–698.

Xuebing, X., C. Jacobsen, N.S. Nielsen, M.T. Heinrich and D. Zhou (2002). Purification and deodorization of structured lipids by short path distillation. Eur. J. Lipid Sci. Thechnol. *104*, 745–755.

Yoon, S.-H., O. Miyawaki, K.-H. Park and K. Nakamura (1996). Transesterification between triolein and ethylbehenate by immobilized lipase in supercritical carbon dioxide. J. Fermentation Bioeng. *82*, 334–340.

Yu, Z.-R., S.S.H. Rizvi and J.A. Zollweg (1992). Enzymatic Esterification of Fatty Acid Mixtures from Milk Fat and Anhydrous Milk Fat with Canola Oil in Supercritical Carbon Dioxide. Biotechnol. Prog. *8*, 508–513.

Zaks, A. and A.M. Klibanov (1988). The effect of water on enzyme action in organic media. J. Biol. Chem. *263*, 8017–8021.

SECTION D

NUCLEIC ACIDS ENZYMES

CHAPTER 21

RESTRICTION AND HOMING ENDONUCLEASES

KRZYSZTOF J. SKOWRONEK AND JANUSZ M. BUJNICKI*

Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, Warsaw, Poland

**iamb@genesilico.pl*

1. INTRODUCTION: HISTORY AND BIOLOGY OF RESTRICTION AND HOMING

1.1. Restriction-Modification Systems

In the early 1950s it was observed that certain bacteria 'restricted' the growth of bacteriophages previously propagated on different strains. In the early 1960s, it was found that the restriction is due to the cleavage of the bacteriophage DNA by endonucleases that are specific for certain sequences and sensitive to covalent modification of bases in the phage genome. Some of these 'restriction endonucleases' (REases) produced discrete DNA fragments upon cleavage, a property very useful for analyzing and recombining the DNA. This opened the door to the rapid development of numerous genetic engineering techniques as well as prompted the search for more REases with novel recognition sequences (reviewed early by Arber and Linn (1969) and in the recent historical perspective (Roberts, 2005)). Thus far, more than 3000 REases with over 250 different sequence specificities have been isolated from various organisms (Roberts *et al*, 2005). The characteristics of REases have been comprehensively reviewed in a book edited by Alfred Pingoud (2004).

Most REases recognize short sequences (4–8 bp) in the DNA and in the presence of Mg^{2+} ions they cleave it in both strands, yielding fragments with a 3′-OH and a 5′-phosphate. Some REases produce 3′ or 5′ overhangs (single stranded extensions at the ends), while others yield blunt ends. Typically, they cleave DNA at the target sites several orders of magnitude more readily than at any alternative sequence. REases are typically accompanied by DNA methyltransferases (MTases) with the same or very similar (sometimes slightly broader) specificity. The MTase uses

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 357–377

S-adenosylmethionine (AdoMet) as a methyl group donor to specifically modify a particular adenine or cytosine in the common target sites in the DNA, thereby preventing the cleavage of these sites by the REase. Methylation provides protection of the 'self' DNA against the destruction. On the other hand, it prompts the destruction of the foreign DNA without the familiar modification pattern, *e.g.* the invading phage DNA (Fig. 1).

Some REases (including those that lead to the discovery of the restriction phenomenon) are not inhibited by methylation – conversely, they attack only methylated DNA (see below for the classification). Genetically linked REase and MTase genes form the so-called restriction-modification (RM) systems, which occur ubiquitously in Bacteria and Archaea and in some viruses that infect lower Eukaryota (Roberts *et al*., 2005). Closely related RM genes appear in phylogenetically distinct

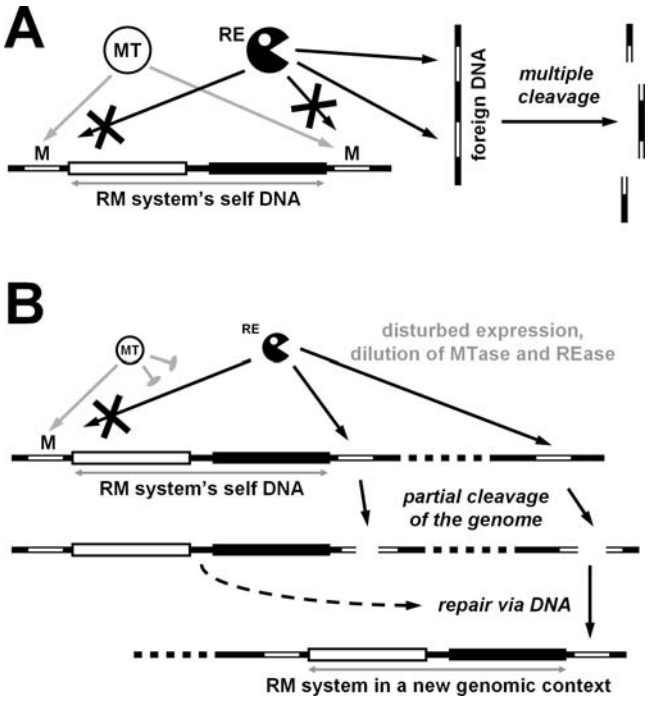


Figure 1. Mechanism of action of RM systems. A) Protection of the self DNA by MTase (white circle, MT), which methylates target sites (indicated in white) and cleavage of the unmethylated sites by REase (black 'pacman', RE). B) Hypothetical mechanism of 'selfish transposition' (Kobayashi, 2001): Disturbed expression of both RM proteins leads to a situation when the diminished MTase activity is insufficient to protect all sites in the genome, and the unmethylated sites are attacked by the remaining REase. If the cleavage is not very extensive, breaks may be repaired and if the entire RM system is reinserted properly at the new location, the expression of MTase may resume to restore the global methylation pattern and to save the cell from death. If the RM system is not properly restored, REase destroys the genome of the host

organisms and in different genomic contexts, and are often associated with mobile genetic elements such as plasmids, viruses, transposons and integrons, which suggests that they have been frequently transmitted by horizontal gene transfer (Jeltsch and Pingoud, 1996; Kobayashi *et al*, 1999). The comparison of closely related bacterial genomes also suggests that, at times, RM genes themselves behave as mobile elements and cause genome rearrangements (Kobayashi, 2001). It is commonly believed that RM systems have been maintained by the prokaryotic cells to defend them from infection by foreign DNAs and to maintain the species identity (review: (Jeltsch, 2003)). An alternative hypothesis suggests that RM systems behave as selfish genetic elements that are maintained because they form parasitic toxin-antitoxin systems whose loss leads to cell death (review: (Kobayashi, 2001). A third hypothesis suggests that RM systems are beneficial 'evolution genes' that serve to increase the genetic diversity, *e.g.* by inducing genome rearrangements (review: (Arber, 2000)). Thus, RM systems can be regarded as intracellular parasites at the verge of mutualism, which can confer advantage in certain contexts (Rocha *et al*, 2001).

1.2. Homing

The discovery of the phenomenon called 'homing' dates back to the early 1970s and the observation of a significant polarity of recombination for markers of an rRNA gene in genetic crosses between yeast mitochondria, which in the 1980s was shown to be due to the action of an intron-encoded 'homing endonuclease' (HEase) (see (Dujon, 2005) for a historical perspective and other articles in a recently published book (Belfort *et al*, 2005) for comprehensive reviews of fundamental studies and applications of HEases). Briefly, homing is a gene conversion process where a mobile sequence is copied and inserted into a cognate allele that lacks this sequence. It is initiated by a site-specific double-strand cut in the target allele, catalyzed by a nuclease encoded within that mobile sequence (reviews: (Belfort *et al*, 1995; Gimble, 2000)). HEases are most often encoded within protein-encoding genes, by introns that self-splice at the RNA level (both group I and group II, and in Archaeal introns) or inteins that self-splice at the protein level, but they can also be free-standing, occurring between genes. Homing is widespread and HEases or their homologs have been found in all phyla.

HEases usually recognize very long DNA sequences (14–40 bp). Unlike REases, they exhibit significant tolerance to substitutions in the target site and effectively recognize whole families of sequences. Nonetheless, the size of the target sites makes them very rare – *e.g.* the historic I-SceI HEase has only one such site in the whole yeast genome of ~13,000 kb. Compared to REases, HEases pose relatively little threat to the integrity of the host DNA, because of the relative rareness of their target sites. Therefore, HEases do not require the assistance of MTases to protect the self DNA by reversible methylation against the multiple rounds of cleavage. Instead, intron- and intein-encoded HEases irreversibly disrupt the target site by inserting the DNA of its mobile element. The insertion sites of free-standing HEases

encoded by T-even phages are, however, spatially separated (up to hundreds of base pairs) from the cleavage sites (Liu *et al*, 2003). Thus, potential cleavage sites for the free-standing HEase exist in both the host (self) and target (non-self) DNA. While the exact mechanism of protection of the self DNA from the cleavage remains unknown, it has been speculated that the nucleotide polymorphism may provide means to discriminate the self and non-self DNA. The cleavage-proficient variants of the self DNA would be counterselected and eliminated, and the resistant variants would be preferentially inherited in the progeny phage (review: (Edgell, 2005)).

Both HEases and REases can be considered as selfish nucleases that promote their own proliferation by cleaving the foreign DNA and thereby inducing DNA repair by recombination, which in turn increases the chance for the duplication of their genes. To minimize the destruction of the self DNA they often resort to modification of the target sites - either irreversibly by disruption of the target DNA in the very process of their own proliferation, or reversibly by enzymatic modification exerted by another enzyme. Interestingly, a close relative of HEases, endonuclease II of phage T4 (Endo II), does not exhibit homing, but is used by the phage to degrade the bacterial DNA, which allows reutilization of the bases for synthesis of the phage DNA (Carlson and Wiberg, 1983). Normal T4 DNA is protected from degradation by modification (hydroxymethylation and glucosylation) of cytosine residues, in striking analogy to the mode of action of RM systems. On the other hand, it was suggested that REases may induce selfish transposition of their RM system, when the balance between restriction enzyme and modification enzyme is somehow disturbed, for example, by the insertion of some genetic element in the neighborhood (Kobayashi, 2001). In this model, the RM gene complex takes advantage of the host's attempt to repair in order to transpose itself to a new locus, similarly to the homing mechanism of type I introns and inteins (see below), but using either ligation of compatible ends generated by cleavage of target sites around the original locus and in the new locus, or (more likely) using non-homologous recombination of exonucleolytically degraded ends (Fig. 1).

2. NOMENCLATURE AND CLASSIFICATION OF REases AND HEases

The rules of nomenclature and functional classification of REases and HEases have been recently updated. Here, we only very briefly review the essentials and the reader should refer to the comprehensive article by Roberts *et al* (Roberts *et al*, 2003). This commonly accepted classification of REases and HEases is robust, purely functional and does not involve any assumptions about the evolutionary relationships between these enzymes. On the other hand, the evolutionary classification of REases and HEases is subject to change as different hypotheses are formulated based on identification of new enzymes and detection of previously

unnoticed relationships. The authors' interpretation of the most recent findings will be reviewed later in this chapter.

2.1. Nomenclature and Classification of REases

The names of REases begin with a three-letter acronym describing the genus (1st letter) and species (2nd and 3rd letter) from which they were isolated and end with the Roman index number. Extra letters and numbers may be added to indicate particular strains or serotypes, *e.g.* EcoRI indicates the first (I) REase isolated from *Escherichia coli R*Y13. Proteins exhibiting REase and MTase actvity may be discriminated by additional letters R. or M. (*e.g.* R.EcoRI and M.EcoRI or RM.EcoKI *etc.*). Putative enzymes (*e.g.* those predicted from sequence analyses but experimentally uncharacterized) or those apparently inactivated, usually include the number of the open reading frame (ORF) and the suffix "P" (*e.g.* HindORF215P).

REases and RM systems have been traditionally classified into three main Types (I, II and III), based mostly on the composition of the proteins and the mode of recognition and cleavage. Type I REases are multisubunit protein complexes that usually comprise separate subunits responsible for different, however coordinated functions: HsdS (specific DNA recognition), HsdM (methylation), and HsdR (ATP-dependent DNA translocation and cleavage). Type I REases cleave the DNA at variable positions away from the target sequence. Type II REases usually act independently of their cognate MTase and will be described in detail below. Type III REases are again multisubunit protein complexes comprising two separate subunits: Mod (specific DNA recognition and methylation) and Res (ATP-dependent DNA translocation and cleavage). Type III REases recognize two copies of the non-palindromic target site that must be in an inverse orientation in the substrate DNA, and they cleave at a fixed position away from one of the two copies of the target sequence. Recently, Type IV has been introduced to describe a heterogeneous group of enzymes with ill-defined sequence specificity and cleaving only modified DNA.

Multiple subtypes of Type II REases (often overlapping with each other) have been introduced to account for different functional peculiarities. Type IIP encompasses all enzymes with "orthodox" features, *i.e.* those recognizing symmetric (palindromic) sequences and cleaving at fixed symmetrical locations within the sequence or immediately adjacent to it. Type IIA enzymes recognize *a*symmetrical sequences. Type IIB REases cut on both sides of the target sequence; their targets may or may not be asymmetric (*i.e.* Type IIP or Type IIA). Type IIC enzymes are unusual in that they include the REase and MTase in the same polypeptide. Type IIE enzymes require an additional copy of the target for cleavage, which itself is not cleaved, but used as an allosteric *e*ffector. Type IIF REases also require two identical sites, but they cleave them both. Type IIG enzymes are a subset of Type IIC that are stimulated or inhibited by AdoMet; their targets may or may not be asymmetric (*i.e.* Type IIP or Type IIA). Type IIH REases exhibit generic features (*e.g.* subunit composition) resembling Type I enzymes, but biochemically behave as Type II. Type IIM REases are not accompanied by MTases, in contrast they

recognize a particular sequence only when it is *m*ethylated and they cleave at a fixed site. Type IIS enzymes are a subset of Type IIA, whose cleavage is shifted outside the target sequence. Type IIT enzymes are composed of heterodimeric subunits.

2.2. Nomenclature and Classification of HEases

Nomenclature of HEases resembles this of REases, with additional prefixes indicating different classes: I- for intron-encoded enzymes (*e.g.* I-CreI), PI- for intein-encoded enzymes (usually with an additional protein splicing activity, *e.g.* PI-SceI), and F- for freestanding enzymes.

HEases encoded by group I introns and Archaeal introns, as well as inteins, exhibit one type of mechanism of homing. They are standalone enzymes that cleave the target site to generate recombinogenic ends, which then engage in a strictly DNA-dependent recombination process that duplicates the intron or intein (including the HEase gene) into the now disrupted target site (Fig. 2). A similar mechanism of action is exhibited by the free-standing HEases, which however (as mentioned earlier) duplicate into a site located at some distance from the cleavage site (review: (Edgell, 2005)). On the other hand, HEases encoded by group II introns utilize a completely different mechanism. Mobility of group-II introns is mediated by a ribonucleoprotein complex comprising the intron RNA and the intron-encoded protein (IEP) with the reverse transcriptase as well as HEase activities. The intron invades the DNA sense strand by splicing using its ribozyme activity, followed by the nicking of the antisense strand by the HEase domain and reverse transcription of the intron RNA by the reverse transcriptase domain of the IEP (review: (Lambowitz and Zimmerly, 2004)).

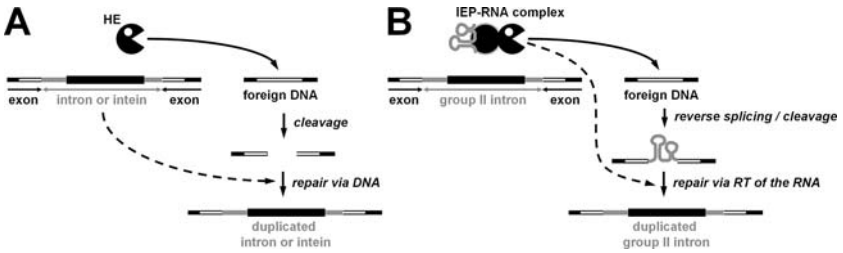


Figure 2. Mobility of HEase-encoding genetic elements. A) The mechanism of homing exhibited by type I and archaeal introns, and inteins (mobile sequence indicated in grey). The HEase (black 'pacman') cleaves the homing site (shown in white) of a cognate intron- or intein-less allele and the break is repaired at the DNA level by gene conversion. A similar mechanism is exhibited by free-standing HEases, only here the gene conversion duplicates the HEase gene at a distance from the cleaveage site. B) The mechanism of retrohoming exhibited by type II introns. The RNA of the IEP-RNA complex invades one strand of the homing site by reverse splicing, and the opposite strand is cleaved by HEase of the IEP. The intron is copied into cDNA by the reverse transcriptase activity of the IEP

3. STRUCTURE AND EVOLUTION OF REases AND HEases

Both REases and HEases are heterogeneous from the structural and evolutionary point of view. All crystal structures of REases solved in the years 1986–2004 revealed the same three-dimensional fold of the PD-(D/E)XK superfamily of nucleases (review: (Pingoud *et al*., 2005)). Therefore, it was often assumed that all members of this group would share a common fold and a similar mechanism of action. However, most of REases show no evident sequence similarity to each other or to any other proteins in the database, which makes sequence-based classification virtually impossible. Even with the availability of state-of-the-art bioinformatic tools, the assignment of REases with unknown structures to structural or evolutionary families remains a challenging task. Thus far, bioinformatic analyses of REase sequences suggested that although indeed many REases belong to the PD-(D/E)XK superfamily, others may belong to other unrelated superfamilies, *e.g.* phospholipase D (PLD), ββαMe, and GIY-YIG ((Aravind *et al*., 2000; Bujnicki *et al*., 2001; Sapranauskas *et al*., 2000), review: (Bujnicki, 2001)). These predictions have been recently supported by experimental analyses (unpublished data; Saravanan *et al*., 2004). Of particular interest is the recent crystallographic structure determination of the Mg^{2+}-independnet, EDTA-resistant nuclease R.BfiI, a relative of phospholipase D (Grazulis *et al*., 2005). It remains to be seen if the so far unassigned REases may belong to some other, structurally and mechanistically different protein superfamilies.

Unlike REases, many HEases show readily detectable sequence similarity to each other. Initially, HEases were thought to belong to four unrelated families: LAGLIDADG, GIY-YIG, HNH and His-Cys box (reviews: (Belfort and Perlman, 1995; Stoddard, 2005)). However, analysis of experimentally determined structures revealed a common active site and suggested that HNH and His-Cys box families are in fact diverged members of the same ββαMe superfamily (Kuhlmann *et al*., 1999). Thus far, the catalytic domains of well-characterized REases and HEases were found to be recruited from five different nuclease superfamilies: LAGLIDADG, GIY-YIG, ββαMe, PD-(D/E)XK, or PLD (Fig. 3), of which GIY-YIG, ββαMe are common to both REases and HEases. Interestingly, the results of preliminary structure predictions for a known cyanobacterial HEase I-SspI suggest it may be a member of the PD-(D/E)XK superfamily (Bujnicki, unpublished data). On the other hand, we have recently identified a large family of PD-(D/E)XK-related proteins in genomes of Cyanobacteria dwelling in fresh waters (Feder and Bujnicki, 2005). In some genomes the number of copies of these putative nucleases reaches 2% of all open reading frames; we predict that at least some of them may be still enzymatically active and engaged in a process similar to intronless homing. This finding suggests that PD-(D/E)XK may be the third superfamily comprising representatives of both REases and HEases.

In addition to the catalytic domains, some REases and HEases possess additional domains, often involved in DNA binding and specific sequence recognition. In particular HEases often feature multiple additional domains tethered to the catalytic domain to provide extensive protein surface for the recognition of their extremely

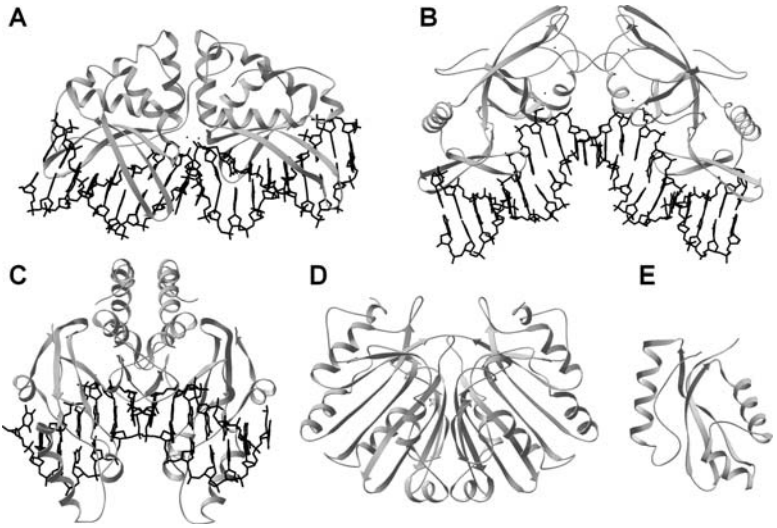


Figure 3. Structures of catalytic domains of representatives of the LAGLIDADG, ββαMe, PD-(D/E)XK, PLD, and GIY-YIG nucleases, respectively: A) I-AniI HEase-DNA complex (a pseudodimer comprising two mutually homologous domains in a single chain), B) I-PpoI HEase-DNA complex (a homodimer), C) PvuII REase-DNA complex (a homodimer), D) BfiI REase (a homodimer with a single active site, C-terminal DNA-binding domain not shown), E) I-TevI HEase (a monomer, C-terminal DNA-binding domain not shown)

long targets. On the other hand, some REases utilize additional DNA-binding domains to either recognize several copies of the same short sequence before the catalysis is triggered (Type IIE) or to position the cleavage site away from the target site (Type IIS). Type IIC REases (including all of Type IIB and IIG and some Type IIH enzymes) employ an inherently non-specific nuclease domain tethered to the sequence-specific DNA MTase module. Interestingly, REases and HEases with unrelated catalytic domains can possess homologous DNA-binding domains and vice versa. For instance, a PD-(D/E)XK-superfamily domain of Type IIE REase R.EcoRII and the PLD-superfamily domain of Type IIS REase R.BfiI are fused to a similar domain from a previously not described superfamily, which also includes B3-like domain of plant transcription factors (Grazulis *et al*., 2005). Likewise, unrelated catalytic domains of HEases I-TevI (GIY-YIG superfamily) and I-HmuI (ββαMe superfamily) are fused to a set of small domains involved in DNA binding, including two that are very similar: a minor groove-binding α-helix and a helix-turn-helix motif (Shen *et al*., 2004; Van Roey *et al*., 2001). Interestingly, a functionally unrelated enzyme, the tRNA splicing endoribonuclease EndA comprises two domains from the LAGLIDADG and PD-(D/E)XK superfamilies, but without the amino acid residues characteristic for the active sites of these nucleases. Instead, the PD-(D/E)XK-like domain evolved an RNase A-like active site in a different part of the structure (Bujnicki and Rychlewski, 2001).

Summarizing, REases and HEases with unrelated catalytic domains have evolved multiple times, recruited various domains to improve the DNA-binding and exchanged them with each other. On the other hand, phylogenetic relationships between REases and HEases in the GIY-YIG and ββαMe superfamilies suggest that these enzymes might have also interconverted their biological function. In particular, some REases might have evolved from free-standing HEases by reducing the length of the original target site, increasing the specificity towards the core base pairs, and associating with the DNA modification enzyme of similar specificity to protect the host genome against the excessive damage (Saravanan *et al*, 2004). Interestingly, it has been shown that REase R.EcoRI can initiate its own homing when placed on an appropriate vector (Eddy and Gold, 1992), which demonstrates the functional overlap between HEases and REases, at least under certain conditions.

4. APPLICATIONS OF REases

REases played key roles in the development of a majority of the techniques used in molecular biology. The number and diversity of different applications of REases make it impossible to provide an exhaustive review on this subject. We will rather try to summarize the common basis of REase uses, providing the subjective selection of methods which exemplify this diversity.

4.1. DNA Sequence Characterization

One group of methods exploits the ability of Type II REases to cleave DNA only within or at a precise distance from the target sequence. Therefore REases can be used to fragment the substrate DNA in a particular pattern that reflects the distribution of target sites in the molecule under study. This feature is used in the ***R**estriction **E**ndonuclease **A**nalysis (REA)* analysis, also called *restriction mapping*. It is by far the most frequently used method of physical characterization or identification of DNA molecules ranging in size from few hundreds to few millions base pairs, including plasmids, DNA viruses and prokaryotic genomes (Allet, 1973; Allet *et al*, 1973; Lee *et al*, 1989; Smith and Condemine, 1990). A restriction map, which can be created either through sequence analysis or through sizing of the fragments generated by the restriction digest of a DNA molecule, defines the arrangement of all fragments generated by a particular enzyme within the whole DNA molecule or its section. Such map then serves as a reference point to interpret the results of experimental fragmentation of the DNA sample under study.

Whole genome restriction mapping is frequently used in epidemiological studies to identify strain of bacterial or viral pathogens or as a first step towards complete physical characterization of genomes by DNA sequencing. The resolution of large fragments used in whole genome analysis requires the application of pulsed-field electrophoresis systems. Frequently, these analyses are limited to a particular region of the genome. In such cases, the products of DNA fragmentation by REase are usually resolved by a regular agarose gel electrophoresis and selective detection of

particular fragments is achieved by hybridization with a radiolabelled DNA (probe) by Southern blotting. This basic technique called ***R**estriction **F**ragment **L**ength **P**olymorphism (RFLP)* analysis found numerous applications in studying eukaryotic genomic DNA, either in basic research or in medical genetic diagnostics, and is widely used to identify different alleles of a particular gene (Owerbach and Nerup, 1982; Wyman and White, 1980).

An interesting variation of RFLP on a genome scale is ***R**estriction **L**andmark **G**enomic **S**canning (RLGS)*. It is practically a genome-wide analysis of restriction pattern obtained in a triple RE digest of the whole genome and fractionation of the products using two-dimensional electrophoresis (Hatada *et al*., 1991; Rush and Plass, 2002). In this technique, DNA fragments obtained after the first REase digest are radiolabelled on both ends, then after the digest with the second REase the mixture of fragments is fractionated in the first dimension, digested in a gel with the third REase, and resolved in the second dimension (Fig. 4). Labelled fragments are detected by autoradiography. By careful selection of REases based on their sensitivity toward DNA methylation, one can also apply this technique to study the methylation state of many regions of the genome in a single experiment (Rush and Plass, 2002).

Another variation of RFLP uses PCR to amplify a selected part of the analyzed genome to target analysis and avoid laborious selective detection by Southern blotting. The amplified region can encompasses the 16S rDNA, the 16S–23S rDNA spacer region and part of the 23S rDNA and such implementation is known as the ***A**mplified **R**ibosomal **D**NA **R**estriction **A**nalysis (ARDRA)* which is frequently used in epidemiological or ecological studies to identify the species and strains of micro-organisms (Deng *et al*., 1992; Gurtler *et al*., 1991; Jayarao *et al*., 1992; Vaneechoutte *et al*., 1992).

The REase activity and PCR amplification is also applied in the ***R**estriction **S**ite **M**utation (RSM)* assay designed to detect rare mutation resulting in loss of a selected restriction site in the background of wild type sequences (Jenkins *et al*., 1990; Parry *et al*., 1990). In this method, PCR is used to selectively amplify the particular underrepresented RFLP. After the sample is digested with a REase, the PCR amplification using primers flanking the site under study is carried out (Fig. 5). Only the DNA that lost the REase target sequence and therefore have sequences complementary to both PCR primers on a single restriction fragment can serve as a PCR template. This assay is frequently used to detect rare mutations in somatic cells at the onset of carcinogenesis (Jenkins, 2004).

4.2. DNA Manipulation

Even in the era of PCR, REases are indispensable in nearly all techniques of *in vitro* DNA manipulations, with molecular cloning (Ausubel *et al*., 2005; Sambrook and Russell, 2000) being probably the most important group of applications. In this case not only the REase specificity but also the unique character of created ends is fully exploited to increase efficiency and selectivity of ligation between

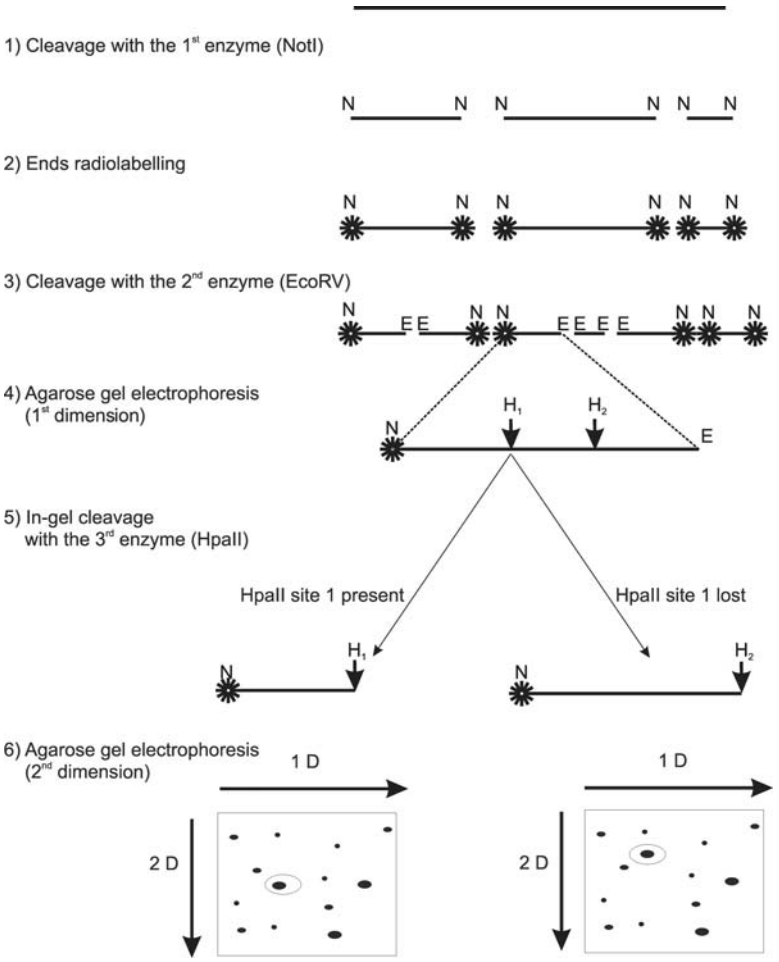


Figure 4. Schematic representation of Restriction Landmark Genomic Scanning technique. The absence of the HpaII site 1 can result from a point mutation or CG methylation (which is inhibitory for HpaII). Those two cases can be discriminated by a similar experiment in which MspI REase will be used instead of HpaII. Both of these enzymes recognize the same sequence (CCGG) but MspI is insensitive to methylation of the 2nd C residue in the CCGG sequence (*i.e.* in the CG dinucleotide)

the insert and the cloning vector. For instance, REase digestion is used as a first step in a classical method of genomic DNA libraries construction. In this case, usually an incomplete digestion is carried out using a REase that recognizes 4 base pairs and is not blocked by DNA methylation characteristic for the source organism and therefore cuts very frequently, nearly randomly throughout the whole genome. This approach has an advantage over other more random enzymatic and physical methods of DNA fragmentation, since created fragments have identical sticky ends,

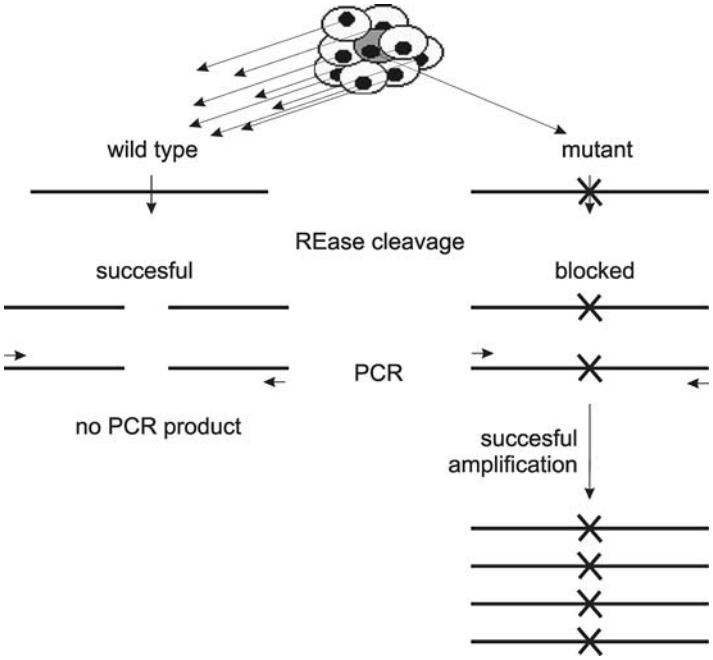


Figure 5. Schematic representation of the Restriction Site Mutation assay, which explains how PCR can amplify a signal from alleles underrepresented in the cell population analyzed (for instance from a few cancer cells in an otherwise healthy tissue)

which can be compatible with the ends of the recipient vector, making the ligation step more efficient. There is a wealth of different variations and tricks in preparation of DNA fragments for cloning or other recombinant DNA techniques, which is beyond the scope of this review (Ausubel *et al*, 2005; Sambrook and Russell, 2000). Another group of REase applications take advantage of the physical properties of the generated ends. Two examples are: specific labeling of one end of the linear double stranded DNA, for instance in DNA footprinting (Brenowitz *et al*, 1986), and generation of unidirectional progressive deletions in Exonuclease III/Mung Bean nuclease system (Henikoff 1984, Henikoff 1987). The DpnI REase that selectively cuts DNA methylated by the Dam MTase (*i.e.* Gm^6ATC) is commonly used in PCR-based site-directed mutagenesis protocols to selectively remove template DNA isolated from *E. coli* leaving product of PCR reaction intact (Weiner *et al*, 1994).

5. APPLICATIONS OF HEases

In contrast to REases, HEases are mainly used *in vivo*. In particular, HEases play an important role in studies of the double strand break (DSB) DNA repair system (Rouet *et al*, 1994). In this case, the HEase gene under control of an inducible promoter is introduced into the cells containing a single target site for this enzyme.

The induction of HEase expression results in formation of the DSB limited to a single locus in a host. Thanks to this selectivity, monitoring of the repair process can be precisely targeted. Since DSBs highly increase the frequency of homologous recombination at a close proximity to the site of lesion, the same setup can be used to controlled and selective gene conversion, which is mechanistically very similar to homing.

HEases are perfect tools for the high-efficiency gene targeting (Donoho *et al*, 1998), which is a crucial part of transgene construction and gene therapy. Site-specific cleavage and formation of double strand breaks (DSBs) in the chromosomal DNA stimulates homologous recombination with a co-trasformed DNA fragment, enabling the introduction of desired mutations by gene replacement (Jasin, 1996). The only obstacle to the universal use of HEases is the limited repertoire of the sequences recognized by the currently known enzymes. Therefore use of HEases for gene targeting must be preceded by introduction of the HEase recognition sequence into an appropriate location in the genome.

The above-mentioned limitation is not present in systems using group II introns, whose specificity is determined by base pairing between short sequence elements in the intron RNA and a > 14-nucleotide region of the DNA target site. Therefore, the target specificity of the HEase can be easily modified by changing the intron RNA (review: (Lambowitz *et al*, 2005)). Group II intron-based integration employs specialized DNA constructs termed targetrons, which contain an inducible promoter to express a cassette encompassing the intron flanked by short exons, and the IEP (*i.e.* the protein including HEase and reverse transcriptase, that is also important for the splicing activity of the intron RNA). However, the IEP gene is removed from the intron and placed downstream of the 3′ exon (Karberg *et al*, 2001; Zhong *et al*., 2003). The IEP expressed from this location efficiently promotes the intron splicing and mobility, but does not accompany the intron when it inserts itself into a new location, thus ensuring its immobility. This system was successfully applied in targeted gene disruption by insertional mutagenesis and in site-specific gene insertion when the gene to be delivered was inserted into the intron sequence. It was demonstrated that group II introns can be engineered to insert into therapeutically relevant DNA target sites in human cells, *e.g.* the HIV-1 provirus and the human gene encoding CCR5, the primary co-receptor for enabling HIV-1 transmission (Guo *et al*, 2000). Engineered group II introns with the inactivated reverse transcriptase to prevent the cDNA synthesis can be also used to introduce targeted DSBs (and thereby stimulate gene replacement by homologous recombination). The DSBs are formed by the reverse splicing and the second strand cleavage, followed by the degradation of the intron RNA by cellular enzymes (review: (Lambowitz *et al*, 2005)).

6. PROTEIN-ENGINEERING OF REASES AND HEASES

A growing number of REase and HEase applications in different aspects of DNA manipulation and characterization leads to an increasing demand for

enzymes recognizing novel sequences. Despite the fact that over 3000 REases with over 250 different sequence specificities have been isolated from various bacterial strains, many possible specificities have not yet been discovered (see *e.g.* http://rebase.neb.com/cgi-bin/classlist). Unfortunately, the success rate for the discovery of novel REases has decreased to a few enzymes per thousands of screened strains over the last several years. The number of experimentally characterized HEases is even smaller, currently less than 200 (however, unlike REases, members of this group of nucleases usually display unique specificities). Thus, there have been substantial efforts to engineer the known REases and HEases to endow them with novel properties, by random mutagenesis as well as rational, *i.e.* structure-guided, design.

Structural data from solved co-crystals of REeases with their cognate DNA sequences indicate that these enzymes employ redundant protein-DNA contacts involving multiple hydrogen bonds to the target sequence, also via water molecules, and combine direct (sequence-based) and indirect (structure-based) read-out for DNA recognition. This saturation and redundancy of protein-DNA contacts allows extraordinary sequence specificity, but also has to be overcome to perform successful protein engineering. REases also often couple recognition with catalysis. Therefore, substitutions of residues involved in base-specific contacts typically lead to null mutants, incapable of either DNA binding or cleavage. Thus far, most attempts to modify the specificity of REases by rational mutagenesis of residues found to be involved in protein-DNA interactions remained unsuccessful (reviews: (Alves and Vennekohl, 2004; Lanio *et al*, 2000)) and only recently some variants with altered specificities have been created by random mutagenesis and screening or selection of mutants. On the other hand, crystal structures of HEases bound to the DNA revealed that these nucleases utilize only a subset of the total potentially available hydrogen-bonding capacity of the DNA sequence. Thus, HEases tolerate significant sequence variation both on the side of the protein and the substrate. This plasticity of the protein-DNA interface greatly facilitates the engineering of HEases with new specificities, which indeed has been more successful than in the case of REases. Here, we will only briefly describe the most successful attempts to engineer the specificity of Type II REases and HEases, with the exception of engineering of HEases encoded by group II introns, which has been covered in the preceding section. For more details the reader should refer to excellent recent reviews (Alves and Vennekohl, 2004; Gimble, 2005).

6.1. Engineering of REases

Based on the analysis of the crystal structure of the R.EcoRV protein-DNA complex (specificity GAT/ATC, cleavage indicated by "/"), Pingoud and co-workers randomized selected residues that could evolve new contacts to the DNA, potentially extending the recognition sequence towards XGAT/ATCY (where X and Y indicate newly recognized, mutually complementary bases). A large-scale *in vitro* screen of mutants led to identification of variants with a 25-fold higher

preference for the A·T pair flanking wild type recognition sequence versus a G·C pair(*i.e.* preferentially cleaving AGAT/ATCT) (Schottler *et al*., 1998; Wenz *et al*., 1998). A similar approach was used to engineer R.BsoBI analogs (specificity C/RCGYG, where R and Y indicate purine and pyrimidine, respectively) to recognize only a subset of the wild type recognition sequences. After randomization of the single residue (D246) selected by the analysis of the R.BsoBI-DNA crystal structure, several variants more stringent then the wild type enzyme were isolated, including the D246A variant showing 70-fold greater binding affinity for the CCCGGG substrate than for CTCGAG (Zhu *et al*., 2003).

Another approach developed by Xu and coworkers employs a high-throughput three-step *in vivo* selection and screening of variants with potentially new specificities (Samuelson and Xu, 2002). The mutant variants are initially selected for survival in a strain expressing a DNA MTase with the desired specificity (which protects only the "new" target sites). Then, to eliminate variants lacking nuclease activity selected variants are screened for the ability to induce DNA cleavage in an indicator strain lacking the DNA MTase. The requirement for a DNA MTase with a "new" specificity limits the applicability of this system. Nonetheless, this approach omits the laborious *in vitro* screening of mutant and allows more extensive randomization of the target REease gene, which is necessary in the case of enzymes with unknown structures. It was used successfully to screen for mutants of R.BstYI (R/GATCY) with specificity increased towards the R.BglII-like sequence (A/GATCT) (Samuelson and Xu, 2002). Recently R. NotI (GC/GGCCGC) was engineered in a two-step approach with the first step similar to the *in vivo* selection described above, leading to isolation of variants with specificity relaxed towards a small set of 8 bp substrates (Samuelson *et al*., 2006). Here, the second step of *in vivo* selection aimed at increasing the enzymatic activity of mutant was based on the cleavage of a conditionally toxic vector carrying the new target sequence.

A conceptually similar screening method called the methylation activity-based selection (MABS) was used to alter the specificity of Type IIG REase R.Eco57I (specificity CTGAAG, cleaves at a short distance away from the recognition site). R.Eco57I comprises the naturally fused, structurally and functionally independent DNA-binding, MTase, and nuclease domains. The nuclease activity of R.Eco57I was inactivated by a single substitution in the active site, followed by error-prone PCR mutagenesis and selection of mutants with the MTase activity specific for new sequences. In a selected variant the nuclease activity was restored by reversion of the initial substitution, resulting in an engineered REase that targets a degenerate sequence CTGRAG (Rimseliene *et al*., 2003). This variant (Eco57MI) is the first commercially available REase with engineered sequence specificity.

6.2. Engineering of HEases

Attempts to engineer sequence specificity of HEases have been reported only recently, but were more successful than in the case of REases. In one of the first attempts to modify the specificity of I-CreI (the most frequent targets of HEase

engineering), the saturating mutagenesis of two amino acid residues critical for DNA recognition was carried out. The mutants were screened *in vivo* for their ability to eliminate a reporter plasmid containing several variants of the altered I-CreI recognition sequence, leading to isolation of enzymes with specificity shifted toward the selected sequence; however, the isolated mutants retained the ability to cleave the wild type target (Seligman *et al*., 2002). In another HEase, PI-SceI, the mutagenesis was targeted to the DNA-binding residues within the protein splicing domain. A bacterial two-hybrid method was used to select PI-SceI variants from a randomized expression library that exhibited altered DNA recognition patterns ranging from partial relaxation of specificity to marked shifts in target site recognition (Gimble *et al*., 2003). A recently developed approach involves a high-throughput selection of HEase variants for the ability to induce DSB-induced homologous recombination in a eukaryotic environment (Chames *et al*., 2005). It has been used to analyze a library of I-CreI mutants with three randomized residues and has led to isolation of hundreds of variants with altered specificities (Arnould *et al*., 2006).

Another method involves altering HEase specificity by domain shuffling. For instance, an active nuclease with modified specificity was created by replacing a DNA recognition region (a subdomain of the protein splicing domain) of PI-SceI, with a homologous region of a related protein PI-CtrIP devoid of the nuclease activity (Steuer *et al*., 2004). Two groups have independently constructed chimeric single-chain HEases by fusing the N-terminal nuclease domain of I-DmoI with a single copy of I-CreI and making different mutations at the domain interface (Chevalier *et al*., 2002; Epinat *et al*., 2003). Expectedly, the resulting chimeric HEases recognize sequences closely resembling the combination of I-DmoI and I-CreI half-sites. The domain shuffling approach can be applied not only to the wild type HEases, but also to the variants engineered by mutagenesis and screening/selection, potentially leading to a plethora of new specificities.

7. CONCLUSIONS AND FUTURE PROSPECTS

While the steadily increasing number of experimentally characterized REases and HEases prompted only modest refinements of their functional classifications compared to those initially proposed, the rapid growth of sequences and structure databases have lead to a complete revision of the picture of sequence-structure-function relationships among these enzymes. It is now commonly accepted that both REases and HEases evolved multiple times, as well as that some families with vastly different sequences, initially believed to be unrelated, in fact share common domains and originated from the same ancestral nuclease superfamilies. Thanks to the combination of crystallographic analyses, sequence comparisons and structural modeling, and large-scale mutational studies, we now begin to understand the basis of sequence specificity, both in terms of the evolution at the level of whole protein families as well as the physical interactions within individual protein-DNA complexes. Even if structure-guided protein engineering still frequently falls short

of providing fully active enzymes with new specificities, progress made recently in this challenging task (especially for HEases) seems to indicate that this goal may soon be reached. The availability of a large number of designed nucleases with new specificities will undoubtedly prompt the development of multiple new applications of these enzymes and improve our abilities to analyze and manipulate the DNA both *in vitro* and *in vivo*.

ACKNOWLEDGEMENTS

We thank Alfred Pingoud for stimulating discussions and Monika Radlinska for critical reading of the manuscript. The authors' research on REases and HEases was funded by the EMBO/HHMI Young Investigator Award to JMB, by the MEiN (grant 3P04A01124), by the E.U. 6th Framework Programme (contract# MRTN-CT-2005-019566), and by the NIH (Fogarty International Center grant R03 TW007163-01).

REFERENCES

Allet, B. (1973) Fragments produced by cleavage of lambda deoxyribonucleic acid with the Hemophilus parainfluenzae restriction enzyme Hpa II. Biochemistry, *12*, 3972–3977.

Allet, B., Jeppesen, P.G., Katagiri, K.J. and Delius, H. (1973) Mapping the DNA fragments produced by cleavage by lambda DNA with endonuclease RI. Nature, *241*, 120–123.

Alves, J. and Vennekohl, P. (2004) Protein engineering of restriction enzymes. In Pingoud, A.M. (ed.), Restriction endonucleases. Springer-Verlag, Berlin, *14*, 393–412.

Aravind, L., Makarova, K.S. and Koonin, E.V. (2000) SURVEY AND SUMMARY: Holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. Nucleic Acids Res, *28*, 3417–3432.

Arber, W. (2000) Genetic variation: molecular mechanisms and impact on microbial evolution. FEMS Microbiol Rev, *24*, 1–7.

Arber, W. and Linn, S. (1969) DNA modification and restriction. Annu.Rev.Biochem., *38*, 467–500.

Arnould, S., Chames, P., Perez, C., Lacroix, E., Duclert, A., Epinat, J.C., Stricher, F., Petit, A.S., Patin, A., Guillier, S., Rolland, S., Prieto, J., Blanco, F.J., Bravo, J., Montoya, G., Serrano, L., Duchateau, P. and Paques, F. (2006) Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. J Mol Biol, *355*, 443–458.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seldman, J.G., Smith, J.A. and Struhl, K. (2005) Current Protocols in Molecular Biology. John Wiley and Sons.

Belfort, M., Derbyshire, V., Stoddard, B.L. and Wood, D.W. (2005) Homing endonucleases and inteins. Springer-Verlag, Berlin.

Belfort, M. and Perlman, P.S. (1995) Mechanisms of intron mobility. J Biol Chem, *270*, 30237–30240.

Belfort, M., Reaban, M.E., Coetzee, T. and Dalgaard, J.Z. (1995) Prokaryotic introns and inteins: a panoply of form and function. J Bacteriol, *177*, 3897–3903.

Brenowitz, M., Senear, D.F., Shea, M.A. and Ackers, G.K. (1986) Quantitative DNase footprint titration: a method for studying protein-DNA interactions. Methods Enzymol, *130*, 132–181.

Bujnicki, J.M. (2001) Understanding the evolution of restriction-modification systems: clues from sequence and structure comparisons. Acta Biochim Pol, *48*, 935–967.

Bujnicki, J.M., Radlinska, M. and Rychlewski, L. (2001) Polyphyletic evolution of type II restriction enzymes revisited: two independent sources of second-hand folds revealed. Trends Biochem Sci, *26*, 9–11.

Bujnicki, J.M. and Rychlewski, L. (2001) Unusual evolutionary history of the tRNA splicing endonuclease EndA: relationship to the LAGLIDADG and PD-(D/E)XK deoxyribonucleases. Protein Sci, *10*, 656–660.

Carlson, K. and Wiberg, J.S. (1983) *In vivo* cleavage of cytosine-containing bacteriophage T4 DNA to genetically distinct, discretely sized fragments. J Virol, *48*, 18–30.

Chames, P., Epinat, J.C., Guillier, S., Patin, A., Lacroix, E. and Paques, F. (2005) *In vivo* selection of engineered homing endonucleases using double-strand break induced homologous recombination. Nucleic Acids Res, *33*, e178.

Chevalier, B.S., Kortemme, T., Chadsey, M.S., Baker, D., Monnat, R.J. and Stoddard, B.L. (2002) Design, activity, and structure of a highly specific artificial endonuclease. Mol Cell, *10*, 895–905.

Deng, S., Hiruki, C., Robertson, J.A. and Stemke, G.W. (1992) Detection by PCR and differentiation by restriction fragment length polymorphism of Acholeplasma, Spiroplasma, Mycoplasma and Ureaplasma, based upon 16S rRNA genes. PCR Methods Appl, *1*, 202–204.

Donoho, G., Jasin, M. and Berg, P. (1998) Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. Mol Cell Biol, *18*, 4070–4078.

Dujon, B. (2005) Homing nucleases and the yeast mitochondrial omega locus - a historical perspective. In Belfort, M., Derbyshire, V., Stoddard, B.L. and Wood, D.W. (eds.), Homing endonucleases and inteins. Springer-Verlag, Berlin, *16*, 11–31.

Eddy, S.R. and Gold, L. (1992) Artificial mobile DNA element constructed from the EcoRI endonuclease gene. Proc.Natl.Acad.Sci.U.S.A., *89*, 1544–1547.

Edgell, D.R. (2005) Free-standing homing endonucleases of T-even phage: freeloaders or functionaries? In Belfort, M., Derbyshire, V., Stoddard, B.L. and Wood, D.W. (eds.), Homing endonucleases and inteins. Springer-Verlag, Berlin, *16*, 147–160.

Epinat, J.C., Arnould, S., Chames, P., Rochaix, P., Desfontaines, D., Puzin, C., Patin, A., Zanghellini, A., Paques, F. and Lacroix, E. (2003) A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. Nucleic Acids Res, *31*, 2952–2962.

Feder, M. and Bujnicki, J.M. (2005) Identification of a new family of putative PD-(D/E)XK nucleases with unusual phylogenomic distribution and a new type of the active site. BMC Genomics, *6*, 21.

Gimble, F.S. (2000) Invasion of a multitude of genetic niches by mobile endonuclease genes. FEMS Microbiol Lett, *185*, 99–107.

Gimble, F.S. (2005) Engineering of homing endonucleases for genomic applications. In Belfort, M., Derbyshire, V., Stoddard, B.L. and Wood, D.W. (eds.), Homing endonucleases and inteins. Springer-Verlag, Berlin, *16*, 177–192.

Gimble, F.S., Moure, C.M. and Posey, K.L. (2003) Assessing the plasticity of DNA target site recognition of the PI-SceI homing endonuclease using a bacterial two-hybrid selection system. J Mol Biol, *334*, 993–1008.

Grazulis, S., Manakova, E., Roessle, M., Bochtler, M., Tamulaitiene, G., Huber, R. and Siksnys, V. (2005) Structure of the metal-independent restriction enzyme BfiI reveals fusion of a specific DNA-binding domain with a nonspecific nuclease. Proc Natl Acad Sci U S A, *102*, 15797–15802.

Guo, H., Karberg, M., Long, M., Jones, J.P., 3rd, Sullenger, B. and Lambowitz, A.M. (2000) Group II introns designed to insert into therapeutically relevant DNA target sites in human cells. Science, *289*, 452–457.

Gurtler, V., Wilson, V.A. and Mayall, B.C. (1991) Classification of medically important clostridia using restriction endonuclease site differences of PCR-amplified 16S rDNA. J Gen Microbiol, *137*, 2673–2679.

Hatada, I., Hayashizaki, Y., Hirotsune, S., Komatsubara, H. and Mukai, T. (1991) A genomic scanning method for higher organisms using restriction sites as landmarks. Proc Natl Acad Sci U S A, *88*, 9523–9527.

Henikoff, S. (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene, *28*, 351–359.

Henikoff, S. (1987) Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol., *155*, 156–165.

Jasin, M. (1996) Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet, *12*, 224–228.

Jayarao, B.M., Dore, J.J., Jr. and Oliver, S.P. (1992) Restriction fragment length polymorphism analysis of 16S ribosomal DNA of Streptococcus and Enterococcus species of bovine origin. J Clin Microbiol, *30*, 2235–2240.

Jeltsch, A. (2003) Maintenance of species identity and controlling speciation of bacteria: a new function for restriction/modification systems? Gene, *317*, 13–16.

Jeltsch, A. and Pingoud, A. (1996) Horizontal gene transfer contributes to the wide distribution and evolution of type II restriction-modification systems. J.Mol.Evol., *42*, 91–96.

Jenkins, G.J. (2004) The restriction site mutation (RSM) method: clinical applications. Mutagenesis, *19*, 3–11.

Jenkins, G.J., Suzen, H.S., Sueiro, R.A. and Parry, J.M. (1999) The restriction site mutation assay: a review of the methodology development and the current status of the technique. Mutagenesis, *14*, 439–448.

Karberg, M., Guo, H., Zhong, J., Coon, R., Perutka, J. and Lambowitz, A.M. (2001) Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. Nat Biotechnol, *19*, 1162–1167.

Kobayashi, I. (2001) Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. Nucleic Acids Res, *29*, 3742–3756.

Kobayashi, I., Nobusato, A., Kobayashi-Takahashi, N. and Uchiyama, I. (1999) Shaping the genome–restriction-modification systems as mobile genetic elements. Curr Opin Genet Dev, *9*, 649–656.

Kuhlmann, U.C., Moore, G.R., James, R., Kleanthous, C. and Hemmings, A.M. (1999) Structural parsimony in endonuclease active sites: should the number of homing endonuclease families be redefined? FEBS Lett, *463*, 1–2.

Lambowitz, A.M., Mohr, G. and Zimmerly, S. (2005) Group II intron homing endonucleases: ribonucleoprotein complexes with programmable target specificity. In Belfort, M., Derbyshire, V., Stoddard, B.L. and Wood, D.W. (eds.), Homing endonucleases and inteins. Springer-Verlag, Berlin, *16*, 121–145.

Lambowitz, A.M. and Zimmerly, S. (2004) Mobile group II introns. Annu Rev Genet, *38*, 1–35.

Lanio, T., Jeltsch, A. and Pingoud, A. (2000) On the possibilities and limitations of rational protein design to expand the specificity of restriction enzymes: a case study employing EcoRV as the target. Protein Eng, *13*, 275–281.

Lee, J.J., Smith, H.O. and Redfield, R.J. (1989) Organization of the Haemophilus influenzae Rd genome. J Bacteriol, *171*, 3016–3024.

Liu, Q., Belle, A., Shub, D.A., Belfort, M. and Edgell, D.R. (2003) SegG endonuclease promotes marker exclusion and mediates co-conversion from a distant cleavage site. J Mol Biol, *334*, 13–23.

Owerbach, D. and Nerup, J. (1982) Restriction fragment length polymorphism of the insulin gene in diabetes mellitus. Diabetes, *31*, 275–277.

Parry, J.M., Shamsher, M. and Skibinski, D.O. (1990) Restriction site mutation analysis, a proposed methodology for the detection and study of DNA base changes following mutagen exposure. Mutagenesis, *5*, 209–212.

Pingoud, A., Fuxreiter, M., Pingoud, V. and Wende, W. (2005) Type II restriction endonucleases: structure and mechanism. Cell Mol Life Sci, *62*, 685–707.

Pingoud, A.M. (2004) Restriction endonucleases. Springer-Verlag, Berlin, Heidelberg.

Rimseliene, R., Maneliene, Z., Lubys, A. and Janulaitis, A. (2003) Engineering of restriction endonucleases: using methylation activity of the bifunctional endonuclease Eco57I to select the mutant with a novel sequence specificity. J Mol Biol, *327*, 383–391.

Roberts, R.J. (2005) How restriction enzymes became the workhorses of molecular biology. Proc Natl Acad Sci U S A, *102*, 5905–5908.

Roberts, R.J., Belfort, M., Bestor, T., Bhagwat, A.S., Bickle, T.A., Bitinaite, J., Blumenthal, R.M., Degtyarev, S., Dryden, D.T., Dybvig, K., Firman, K., Gromova, E.S., Gumport, R.I., Halford, S.E., Hattman, S., Heitman, J., Hornby, D.P., Janulaitis, A., Jeltsch, A., Josephsen, J., Kiss, A., Klaenhammer, T.R., Kobayashi, I., Kong, H., Kruger, D.H., Lacks, S., Marinus, M.G., Miyahara, M., Morgan, R.D., Murray, N.E., Nagaraja, V., Piekarowicz, A., Pingoud, A., Raleigh, E., Rao,

D.N., Reich, N., Repin, V.E., Selker, E.U., Shaw, P.C., Stein, D.C., Stoddard, B.L., Szybalski, W., Trautner, T.A., Van Etten, J.L., Vitor, J.M., Wilson, G.G. and Xu, S.Y. (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. Nucleic Acids Res, *31*, 1805–1812.

Roberts, R.J., Vincze, T., Posfai, J. and Macelis, D. (2005) REBASE–restriction enzymes and DNA methyltransferases. Nucleic Acids Res, *33*, D230–232.

Rocha, E.P., Danchin, A. and Viari, A. (2001) Evolutionary role of restriction/modification systems as revealed by comparative genome analysis. Genome Res, *11*, 946–958.

Rouet, P., Smih, F. and Jasin, M. (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol, *14*, 8096–8106.

Rush, L.J. and Plass, C. (2002) Restriction landmark genomic scanning for DNA methylation in cancer: past, present, and future applications. Anal Biochem, *307*, 191–201.

Sambrook, J. and Russell, D.W. (2000) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Samuelson, J.C., Morgan, R.D., Benner, J.S., Claus, T.E., Packard, S.L. and Xu, S.Y. (2006) Engineering a rare-cutting restriction enzyme: genetic screening and selection of NotI variants. Nucleic Acids Res, *34*, 796–805.

Samuelson, J.C. and Xu, S.Y. (2002) Directed evolution of restriction endonuclease BstYI to achieve increased substrate specificity. J Mol Biol, *319*, 673–683.

Sapranauskas, R., Sasnauskas, G., Lagunavicius, A., Vilkaitis, G., Lubys, A. and Siksnys, V. (2000) Novel subtype of type IIs restriction enzymes. BfiI endonuclease exhibits similarities to the EDTA-resistant nuclease Nuc of *Salmonella typhimurium*. J Biol Chem, *275*, 30878–30885.

Saravanan, M., Bujnicki, J.M., Cymerman, I.A., Rao, D.N. and Nagaraja, V. (2004) Type II restriction endonuclease R.KpnI is a member of the HNH nuclease superfamily. Nucleic Acids Res, *32*, 6129–6135.

Schottler, S., Wenz, C., Lanio, T., Jeltsch, A. and Pingoud, A. (1998) Protein engineering of the restriction endonuclease EcoRV–structure-guided design of enzyme variants that recognize the base pairs flanking the recognition site. Eur J Biochem, *258*, 184–191.

Seligman, L.M., Chisholm, K.M., Chevalier, B.S., Chadsey, M.S., Edwards, S.T., Savage, J.H. and Veillet, A.L. (2002) Mutations altering the cleavage specificity of a homing endonuclease. Nucleic Acids Res, *30*, 3870–3879.

Shen, B.W., Landthaler, M., Shub, D.A. and Stoddard, B.L. (2004) DNA binding and cleavage by the HNH homing endonuclease I-HmuI. J Mol Biol, *342*, 43–56.

Smith, C.L. and Condemine, G. (1990) New approaches for physical mapping of small genomes. J Bacteriol, *172*, 1167–1172.

Steuer, S., Pingoud, V., Pingoud, A. and Wende, W. (2004) Chimeras of the homing endonuclease PI-SceI and the homologous Candida tropicalis intein: a study to explore the possibility of exchanging DNA-binding modules to obtain highly specific endonucleases with altered specificity. Chembiochem, *5*, 206–213.

Stoddard, B.L. (2005) Homing endonuclease structure and function. Q Rev Biophys, 1–47.

Van Roey, P., Waddling, C.A., Fox, K.M., Belfort, M. and Derbyshire, V. (2001) Intertwined structure of the DNA-binding domain of intron endonuclease I-TevI with its substrate. Embo J, *20*, 3631–3637.

Vaneechoutte, M., Rossau, R., De Vos, P., Gillis, M., Janssens, D., Paepe, N., De Rouck, A., Fiers, T., Claeys, G. and Kersters, K. (1992) Rapid identification of bacteria of the Comamonadaceae with amplified ribosomal DNA-restriction analysis (ARDRA). FEMS Microbiol Lett, *72*, 227–233.

Weiner, M.P., Costa, G.L., Schoettlin, W., Cline, J., Mathur, E. and Bauer, J.C. (1994) Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. Gene, *151*, 119–123.

Wenz, C., Hahn, M. and Pingoud, A. (1998) Engineering of variants of the restriction endonuclease EcoRV that depend in their cleavage activity on the flexibility of sequences flanking the recognition site. Biochemistry, *37*, 2234–2242.

Wyman, A.R. and White, R. (1980) A highly polymorphic locus in human DNA. Proc Natl Acad Sci U S A, *77*, 6754–6758.

Zhong, J., Karberg, M. and Lambowitz, A.M. (2003) Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. Nucleic Acids Res, *31*, 1656–1664.

Zhu, Z., Zhou, J., Friedman, A.M. and Xu, S.Y. (2003) Isolation of BsoBI restriction endonuclease variants with altered substrate specificity. J Mol Biol, *330*, 359–372.

CHAPTER 22

DNA POLYMERASES FOR PCR APPLICATIONS

RÉGEN DROUIN*, WALID DRIDI AND OUMAR SAMASSEKOU

Service of Genetics, Department of Pediatrics, Faculty of Medicine & Health Sciences, Université de Sherbrooke, Sherbrooke, Canada

**regen.drouin@usherbrooke.ca*

1. INTRODUCTION

In 1956 Arthur Kornberg discovered DNA polymerase I of *Escherichia coli,* the first enzyme implicated in the replication process to be identified (Kornberg, 1957). The principal function of DNA polymerases is to copy DNA using one of its strands as a template and employing small fragments of DNA or RNA as primers for elongation from the 5' end to the 3'-OH end. In addition to this function, DNA polymerases are also involved in the maintenance of genome integrity during DNA replication, DNA repair, homologous recombination, sister chromatid cohesion, cell cycle checkpoints and the development of the immune system. The dysfunction of DNA polymerases can cause diseases such as external ophthalmoplegia (Copeland and Longley, 2003; Ponamarev *et al.*, 2002), Xeroderma pigmentosum (Kannouche and Stary, 2003) and the process of tumourigenesis (Loeb, 2001).

DNA polymerases have revolutionised molecular biology with their ability to amplify small amounts of DNA *in vitro*. Over the last 20 years their use in the polymerase chain reaction (PCR) has overcome a major limiting factor in daily medicine *i.e.* the quantitative problem of the small amounts of DNA available for testing. These small amounts of DNA can be a single gene, or just part thereof. Distinct from replication in living organisms, the PCR process can amplify only relatively short DNA fragments, usually up to 10 kb. However, with the advent of recombinant polymerases amplified fragments up to 70 kb long can be achieved (Blanco *et al.*, 1989). Commercial DNA polymerases come from various species, and they differ in their structures, their catalytic properties such as processivity, their fidelity of proofreading and their rate of extension of the DNA strand (Bohlke *et al.*, 2002).

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 379–401

The PCR technique was developed in 1985 by K. Mullis (Cetus Corporation) (Mullis *et al*, 1986), who received the Nobel Prize in Chemistry in 1993 for this discovery. The original PCR procedure used heating at 94°C in order to denature DNA - a temperature that destroyed the DNA polymerase used, hence the enzyme had to be replenished after the heating stage of each cycle. The original PCR technique was thus very inefficient and labour-intensive since it required much time, large amounts of DNA polymerase and continual attention throughout the process. The idea of using the thermoresistant DNA polymerase from *Thermus aquaticus*, a thermophilic bacterium described in 1969 (Brock and Freeze, 1969), resolved the problem and gave birth to the modern PCR technique. Within just a few years of its invention PCR was being widely used for various purposes, making it one of the most spectacular developments in the history of molecular biology.

Despite the simplicity of its principle and realization, the PCR technique hides numerous snares that can compromise the value of the results obtained. Effective use of PCR requires good laboratory organization, adequate knowledge of each PCR step, experience and the choice of an appropriate DNA polymerase for the task intended.

Polymerases commonly used for PCR are obtained from various thermophilic micro-organisms: *Thermus aquaticus* (Taq), *Pyrococcus furiosus* (Pfu polymerase), *Thermococcus litoralis* (Wind or Tli polymerase or Vent polymerase) and *Thermus thermophilus* (Tth polymerase). The use of a specific polymerase depends on the type of PCR being undertaken and the nature and the size of template. Polymerase mixtures have been developed to overcome difficulties associated with the use of single polymerases in certain types of PCR. Nowadays, the PCR technique is used on a daily basis in medical and biological research laboratories for numerous tasks such as the detection of genetic diseases, the identification of genetic fingerprints, gene cloning, paternity testing and the diagnosis of infectious diseases by detection of bacteria or viruses (particularly AIDS).

2. STRUCTURE OF DNA POLYMERASES

A number of different DNA polymerases have been discovered recently: five have been recognized in *Escherichia coli*, nine in *Saccharomyces cerevisiae* and sixteen in humans (Table 1) (Bebenek and Kunkel, 2004). Variations in the amino acid sequences in the catalytic subunits of the DNA polymerases have been used to define five families for these enzymes: A, B, C, X and Y (Table 1) (Braithwaite and Ito, 1993; Burgers *et al*, 2001). Nevertheless, all DNA polymerases share a common structural conformation akin to the human right hand comprised of three distinct domains designated palm, thumb and fingers (Fig. 1) (Franklin *et al*, 2001; Lewin, 2004). Conserved sequence motifs in the palm provide the catalytic active site; correct positioning of the template at the active site is effected by the fingers; the thumb is involved in binding the DNA and plays a role in processivity of the enzyme (Steitz, 1998). Conserved regions of these three domains coincide at the catalytic site.

Table 1. DNA polymerases in *Escherichia coli, Saccharomyces cerevisiae* and Humans*

Family	Name	Bacterial gene	Human gene	Yeast gene	Mol. Wt. (kDa)[a]	3'Exo activity	Other activities
A	Ec Pol I	*Pol A*			103	+	5' Exonuclease
	γ (gamma)		*POLG*	*MIPI*	140	+	dRP lyase
	θ (theta)		*POLQ*	-	290	-	ATPase, helicase
	ν (nu)		*POLN*	-	100	-	
B	Ec Pol II	*Pol B*			89	+	
	α (alpha)		*POLA*	*POL1(CDC17)*	165	-	Primase
	δ (delta)		*POLD1*	*POL3 (CDC3)*	125	+	
	ε (epsilon)		*POLE*	*POL2*	225	+	
	ζ (zeta)		*POLZ (REV3)*	*REV3*	353	-	
C	Ec Pol III	*dnaE*			130	(separate subunit)	
X	β (beta)		*POLB*	-	39	-	dRP lyase, AP lyase
	λ (lambda)		*POLL*	*POL 4 (POLX)*	66	-	DRP lyase, TdT
	μ (mu)		*POLM*	-	55	-	TdT
	TdT		*TdT*		56	-	
	σ (sigma)		*POLS (TRF4-1)*	*TRF4*	60	-	
Y	Ec Pol IV	*dinB*			40	-	
	Ec Pol V	*umuC*			46		
	η (eta)		*POLH (RAD30A, XPV)*	*RAD30*	78	-	
	ι (iota)		*POLI (RAD30B)*	-	80	-	dRP lyase
	κ (kappa)		*POLK 9DINB)*	-	76	-	
	Rev1		*REV1*	*REV1*	138	-	

* From Bebenek and Kunkel, 2004, p 138.

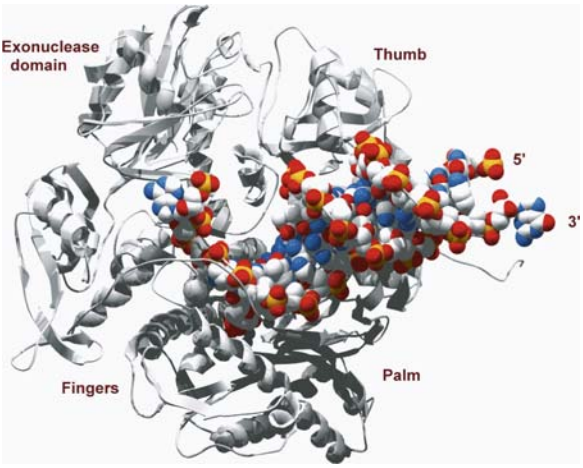


Figure 1. Structure of a replicative DNA polymerase (Franklin *et al*., 2001, p 660)

Six highly conserved regions termed I–VI have been identified among eukaryotic, prokaryotic and viral polymerases. Region IV is the most N-terminally located, followed by regions II, VI, III, I and V. Region I is located in the palm and contains two conserved aspartic acid residues. The other invariant aspartic acid occurs in region II and is located at the tip of a β-sheet that is part of the palm subdomain. Included in this region is the highly conserved SLYPS-II region, which is important for deoxynucleoside triphosphate (dNTP) binding. Other residues important for dNTP binding are in region III. Region IV is located at the N-terminus that is part of the 3' to 5' exonuclease active site. The other two conserved regions, V and VI, are located in the thumb and finger subdomains, respectively (Hubscher *et al*., 2002).

The phosphoryl transfer that takes place during polymerisation is catalysed by a two-metal-ion mechanism which also plays an important role in the exonuclease activity (Sawaya *et al*., 1994; Steitz, 1998). Two Mg^{2+} ions form a pentacoordinated transition state with the phosphate groups of the incoming nucleotide via interactions with conserved carboxylate residues in regions I and II. In addition to this feature, the finger subdomain rotates toward the palm to move from an 'open' to a 'closed' conformation forming the binding groove for the incoming dNTP. The thumb domain also rotates toward the palm. The resulting closed conformation allows interactions between the conserved residues of the fingers with the dNTP binding site and the exonuclease domain (Franklin *et al*., 2001; Lewin, 2004).

Polymerases α, δ, ε and ζ (family B) and γ (family A) are well characterized in eukaryotes. They are heteromultimers and are composed of a large subunit and a variety of smaller subunits. The latter have been implicated in the stabilization of the catalytic subunit, establishment of protein-protein interactions, cell-cycle regulation and also checkpoint function (Hubscher *et al*., 2002).

Polymerase α is a heterotetrameric enzyme. The N-terminal domain seems to be dispensable for catalytic activity and the assembly of the tetrameric complex. A central domain contains all the conserved regions responsible for DNA binding, dNTP binding and phosphoryl transfer. The C-terminal domain is not essential for catalysis but is necessary for the interaction with the other subunits. Pol α might be involved in cell regulatory functions (Hubscher *et al*., 2002).

Polymerase β is the smallest eukaryotic polymerase and consists of two domains. The N-terminal domain performs the 5'-deoxyribose phosphatase activity (to remove the 5'-deoxyribose phosphate) and single-stranded DNA binding, whereas the large domain carries out the polymerase activity. Polymerase β is able to fill short gaps in a distributive way and these gaps contain a 5'-phosphate (Beard and Wilson, 2006).

The N-terminal part of polymerase δ contains three regions of high homology: a nuclear targeting signal and nuclear targeting regions 1 and 2. The C-terminal part has three highly similar regions termed CT-1 to 3 and a zinc-finger domain. Polymerase δ might be involved in double-strand break repair (Hubscher *et al*., 2002).

Polymerase γ is a heterodimeric protein composed of a large subunit, which is responsible for the catalytic activities (DNA polymerase and both the 5' to 3' and 3' to 5' exonuclease activities - Graves *et al*., 1998), and a small accessory subunit. The small subunit dimerizes with the large one and binds DNA. Pol γ is one of the key players in mitochondrial DNA repair (Vanderstraeten *et al*., 1998).

The C-terminal region of polymerase ε contains a highly acidic region and a zinc-finger domain. Polymerase ε is implicated in DNA replication in budding yeast and DNA repair in human cells (D'Urso and Nurse, 1997; Pospiech *et al*., 1999).

3. CATALYTIC PROPERTIES OF DNA POLYMERASES

DNA polymerases have different enzymatic activities:

5' to 3' DNA polymerase activity: DNA polymerases catalyse the linkage of dATP, dCTP, dGTP and dTTP in a specific order, using single-stranded DNA as a template such that the newly polymerised molecule is complementary to the template. DNA polymerases synthesize DNA in the 5' to 3' direction. They are unable to begin a new chain *de novo* and require a pre-existing 3'-OH group in the form of a primer to which the first nucleotide of a new chain is added. Primers are RNA in most organisms but can be DNA in some; in the case of certain viruses the primer is a protein the presents a nucleotide to the polymerase (Lewin, 2004). To complete DNA strand synthesis RNA is removed from the fragments, a DNA polymerase fills the gaps and the nicks remaining are ligated.

3' to 5' exonuclease, proofreading activity: error correction by proofreading is a property of some DNA polymerases. This process corrects mistakes in newly synthesized DNA. When an incorrect base pair is recognized DNA polymerase reverses its direction by one or more base pairs of DNA. The 3' to 5' exonuclease activity of the enzyme allows the incorrect base to be excised. Following proofreading, the polymerase re-inserts the correct base and replication continues.

Generally, DNA polymerases lacking 3' to 5' exonuclease activity have higher error rates than the polymerases that possess it (Table 2) and tend to pause after inserting incorrect bases.

5' to 3' exonuclease, nick translation activity: this enzymatic activity plays an essential role in DNA replication by removing RNA primers.

Terminal transferase activity: causes the addition of a single nucleotide (generally adenine) to the 3' end of PCR products.

4. TYPES OF DNA POLYMERASES COMMERCIALLY AVAILABLE

DNA polymerase I. This is a bacterial enzyme that plays an important role in DNA repair and a secondary role in replication (Camps and Loeb, 2004). The commercial form is extracted from *E. coli*. Its 5' to 3' DNA polymerase activity requires a template; it also has 3' to 5' and 5' to 3' exonuclease activity. The DNA synthesized is strictly complementary to the template. This enzyme is used to synthesize DNA from a single-strand DNA template at 37°C. Its major applications are: the determination of a DNA sequence using an enzymatic sequencing method, the synthesis of radioactive probes, transformation of a staggered end into a double-stranded blunt end, and the construction of vectors from a single DNA strand.

The Klenow Fragment of DNA polymerase I. The Klenow fragment is a fragment of DNA polymerase I obtained by limited proteolysis. The 5' to 3' exonuclease activity is removed while preserving the 5' to 3' polymerase and the 3' to 5' exonuclease activities (Sanger *et al*., 1977).

T4 DNA polymerase. Isolated from bacteriophage T4 of *E. coli*. It has the same enzymatic activities as the Klenow fragment and the same biological uses (Kaplan and Delpech, 1994).

Taq DNA Polymerase. This is a thermostable enzyme isolated from *Thermus aquaticus* which is used for PCR amplification of DNA fragments up to 5 kb in length, as well as DNA labelling and sequencing. It is ideal for TA (T for 'T vector' (thymidine); A for adenosine) cloning (Zhou and Gomez-Sanchez, 2000). This procedure exploits the enzyme's terminal transferase activity. Taq polymerase has a non-template dependent activity which adds a single adenosine to the 3' ends of double-stranded DNA molecules. Thus most molecules PCR amplified by Taq polymerase possess single 3'-A overhangs. The use of a linearized 'T-vector' which has single 3'-T overhangs allows direct, high-efficiency cloning of PCR products facilitated by the complementarities between the PCR product 3'-A overhangs and the 3'-T overhangs of the vector (Zhou and Gomez-Sanchez, 2000).

Terminal transferase. This is a mammalian enzyme expressed in lymphocytes. The enzyme purchased commercially is usually produced by expression of the bovine gene DNTT in *E. coli*. It catalyses deoxynucleotide addition to a free 3'-OH end without the need for a template. The choice of deoxynucleotide added is made randomly. The base composition of the synthesized polydeoxynucleotide depends on the base concentrations in the incubation medium. Terminal transferase

Table 2. DNA polymerases characteristics

DNA polymerase	Error rate (error/bp)	3' to 5' exonuclease	Terminal transferase activity	Half-life	Price per U* in US$	Applications	Companies
Taq	1.1×10^{-4}	-	+	60 min at 94 °C	0.46$	¬ TA cloning ¬ DNA labelling ¬ Sequencing	QIAGEN
Deep Vent®	4.5×10^{-5}	+	-	23 h at 95 °C 8 h at 100 °C	0.42$	¬ High-fidelity PCR ¬ Primer-extension ¬ Cloning	NEW ENGLAND *BioLabs*
Pfu DNA polymerase	1.6×10^{-6}	+	-	19 h at 95 °C	1.74$	¬ High-fidelity PCR ¬ Primer-extension ¬ Cloning ¬ Blunt-end amplification product generation	STRATAGENE
Herculase® Enhanced	2.8×10^{-6}	+	+	NF	1.28$	¬ GC rich fragment	STRATAGENE
Phusion™	4.4×10^{-7}	+	-	> 6 h at 96 °C	1.45$	¬ High-fidelity PCR ¬ Fast long amplicon	*FINNZYMES*
T4 DNA polymerase	1×10^{-6}	+	-	NF	0.38$	¬ Form blunt ends ¬ Probe labelling	NEW ENGLAND *BioLabs*
Klenow Fragment	18×10^{-6}	+	NF	Heat inactivated: 70 °C for 10 minutes	0.27$	¬ Labelling recessed 3' end of double stranded DNA ¬ Random-priming DNA labelling	*FERMENTAS*

(*Continued*)

Table 2. (*Continued*)

DNA polymerase	Error rate (error/bp)	3' to 5' exonuclease	Terminal transferase activity	Half-life	Price per U* in US$	Applications	Companies
Terminal transferase	NA	NA	+	Heat inactivated: 75 °C for 20 minutes	0.11$	¬ Tailing ¬ Labelling the DNA 3-OH ends	NEW ENGLAND *BioLabs*
Sequenase™ version 2.0	3.4×10^{-5}	-	+	NF	0.70$	¬ Sequencing	USB
DNA polymerase I	NF	+	NF	Heat inactivated: 75 °C for 10 minutes	0.15$	¬ Sequencing ¬ Transforming a single blunt end into a double stranded blunt end ¬ Vector construction	USB
rTth DNA polymerase XL		+	NF	NF	0.58$	¬ Amplify long fragments > 40 kb	Applied Biosystems
Isis proofreading DNA polymerase	0.66×10^{-6}	+	-	18 h at 95 °C and 5 h at 100 °C	1.40$	¬ GC rich fragments, or contain secondary structure ¬ Cloning ¬ Characterization of rare mutations in tissue	Krackeler Scientific Inc
rBst DNA polymerase	NF	+	NF	Optimal activity at 65 °C	0.50$	¬ GC rich fragments, or contain secondary structure	EPICENTRE Biotech-nologies

phi29 DNA polymerase	NF	+	NF	Heat inactivated: 65 °C for 10 minutes	0.23$	¬ Rolling circle replication ¬ Multiple displacement amplification ¬ Unbiased amplification of whole genome ¬ Amplify long fragments > 70 kb	NEW ENGLAND *BioLabs*
SurePRIME™DNA polymerase	NF	NF	NF	> 40 min at 95 °C	0.80$	¬ High-fidelity PCR, even when using non-optimised primers ¬ Hot Start PCR	Krackeler Scientific, Inc.
BioTHERM™DNA polymerase	NF	NF	+	NF	0.13$	¬ PCR with low-abundance template ¬ TA cloning	GENECRAFT
SpeedSTAR™ HS DNA polymerase	NF	-	+	NF	0.87$	¬ High speed amplification	Takara
MTP™ Taq DNA polymerase	NF	-	NF	NF	0.90$	¬ Detecting and identifying bacterial DNA	SIGMA-ALDRICH
KOD "Hot Start" DNA Polymerase	NF	+	-	NF	0.90$	¬ High speed amplification ¬ Hot Start PCR ¬ High-fidelity PCR	Novagen

* U: One unit is defined as the amount of enzyme required to incorporate 10 nmoles of dNTPs into acid precipitable form after 30 minutes at 70°C
NF: Not Found
NA: Non applicable

is an example of a DNA polymerase that does not require a primer. Cobalt is a necessary cofactor for activity of this enzyme. The enzyme lacks 3' to 5' and 5' to 3' exonuclease activities. It is used to generate DNA blunt ends and for labelling of DNA 3' ends (Chang and Bollum, 1986).

Pfu DNA polymerase. This polymerase derives from the hyperthermophilic archae *Pyrococcus furiosus.* It exhibits a long half-life and proofreading properties comparable to those of other thermostable polymerases. It has 3' to 5' exonuclease activity and a high proofreading efficiency and lacks 5' to 3' exonuclease and terminal transferase activities. It is used for high-fidelity PCR and primer-extension reactions (Angers *et al.*, 2001), PCR cloning and the generation of blunt-end amplification products.

Sequenase. The sequenases are a family of polymerases derived from the DNA polymerase of bacteriophage T7. They are dimers of the gene 5 protein of the phage and a protein produced by its host cell (thioredoxin). They are the fastest of all DNA polymerases and are used in sequencing with dideoxyribonucleotides (the Sanger method).

Deep Vent. Isolated from *Thermococcus litoralis*; also known as Tli polymerase (Jannasch *et al.*, 1992). Used for high-fidelity PCR and primer-extension reactions.

Herculase Enhanced. A mixture of the high fidelity Pfu polymerase, Taq polymerase and ArchaeMaxx enhancing factor. Used to amplify long and GC-rich DNA fragments.

Phusion. A novel enzyme with extreme fidelity and high speed, used for PCR cloning. It allows high product yields with minimal enzyme concentrations.

rTth DNA polymerase. A recombinant enzyme from *Thermus thermophilus* used to amplify DNA fragments of more than 40 kb (Fromenty *et al.*, 2000).

rBst DNA polymerase. The product of the DNA *pol* I gene of the thermophilic bacterium *Bacillus stearothermophilus* (*Bst*) produced in *E. coli*. It can synthesize DNA regions of high GC content where other non-thermostable DNA polymerases may fail. It is used for replicating difficult templates such as those that contain hairpin structures (Ye and Hong, 1987).

phi29 DNA polymerase. Obtained from *Bacillus subtilis* phage phi29. This enzyme acts preferentially on single-stranded DNA and can synthesize stretches of more than 70 kb. However, its half-life is only 10 minutes at 65°C (Blanco *et al.*, 1989). It is a very accurate polymerase and can yield large amounts of amplified DNA even from small amounts of template. It is suitable for rolling circle amplification, multiple displacement amplification, unbiased whole genome amplification, protein-primed DNA amplification and *in situ* genotyping with padlock probes.

Isis proofreading DNA polymerase. From *Pyrococcus abyssi*. One of the most thermostable proofreading polymerases available, permitting highly accurate DNA synthesis (Gueguen *et al.*, 2001).

SurePRIME™ DNA polymerase. This enzyme is a highly purified form of recombinant Taq polymerase that has been chemically modified by the addition of heat-labile blocking groups to specific amino acid residues. Prior to the PCR

reaction, the enzyme is in an inactive state. It is incapable of extending primer-dimers or mis-annealed primer-template species that form below the specific annealing temperature. Primer-dimers can occur when two primers, or parts of them, are complementary and hybridise to each other. The 95°C incubation step therefore serves to activate the enzyme and also ensure a completely 'clean' initial PCR cycle. This procedure is called 'Hot Start' PCR.

BioTherm*[tm] *DNA polymerase. A thermostable DNA polymerase purified from *Thermus aquaticus*. It is used for PCR primed off a low-abundance template (less than 100 DNA molecules).

SpeedSTAR*[TM] *HS DNA polymerase. Allows very fast extension.

MTP*[TM] *Taq DNA Polymerase. A recombinant Taq DNA polymerase specifically useful for applications involving the detection and identification of bacterial DNA.

KOD DNA Polymerase. Isolated from the extreme thermophile *Thermococcus kodakaraensis* KOD1. It is commercially available in three different versions: KOD HiFi is able to amplify targets up to 6 kb. KOD Hot Start is KOD HiFi DNA polymerase premixed with two monoclonal antibodies that inhibit the DNA polymerase and 3' to 5' exonuclease activities at ambient temperatures. It is able to amplify longer targets than KOD HiFi alone (up to 21 kb with plasmid DNA template), in addition to having the advantage of room temperature set up and fewer mis-priming problems. It can synthesize DNA in regions containing particular secondary structures or a high GC content and generates blunt-ended PCR products suitable for cloning. KOD XL DNA polymerase is an optimised blend of KOD HiFi DNA polymerase and a mutant form of KOD HiFi that is deficient in 3' to 5' exonuclease activity. It is designed for the amplification of longer (up to 30 kb) and more complex GC rich targets. KOD XL DNA polymerase generates a mixture of PCR products with blunt and 3'-dA overhangs.

5. THE PCR TECHNIQUE

DNA amplification by PCR requires a DNA template which contains the targeted region. Best results are obtained by using well-purified DNA lacking contamination by RNA or protein. Two primers that determine the beginning and end of the region to be amplified are required. The choice of primers is critical, and several factors must be carefully considered, *e.g.* the melting temperature (Tm) which is defined as the temperature at which half of the primer binding sites are occupied, GC content, the presence of secondary structures. The sequence of the primers should be such as not to permit the formation of hairpins or hybrids between them, and programs are available to help in their design *e.g.* Oligo and Gene Jockey. The choice of DNA polymerase depends on the specific application (see section 4). Deoxynucleotide triphosphates (dATP, dCTP, dGTP, dTTP) are the monomers from which the DNA polymerase synthesizes the new DNA. The buffer used for a PCR reaction maintains the pH at the optimal level for the polymerase. Bivalent cations (Mg^{2+}) and monovalent cations (K^+, Na^+ or NH_4^+) are necessary to neutralize the

negative charges of the phosphate groups of the DNA and to stabilize DNA/DNA hybrids. The PCR reaction is carried out in a thermal cycler which successively heats and cools the reaction tubes to the precise temperatures and for the specific periods required for each step of the reaction. The PCR technique usually consists of three principal steps - denaturation, primer annealing and extension - which are repeated for 20 to 30 cycles (Fig. 2).

PCR is a powerful tool but errors and mistakes can easily occur. The polymerase reaction is very sensitive to different variables. Divalent cations, especially Mg^{2+}, play an important role in nucleotide stability and affect the polymerisation activity

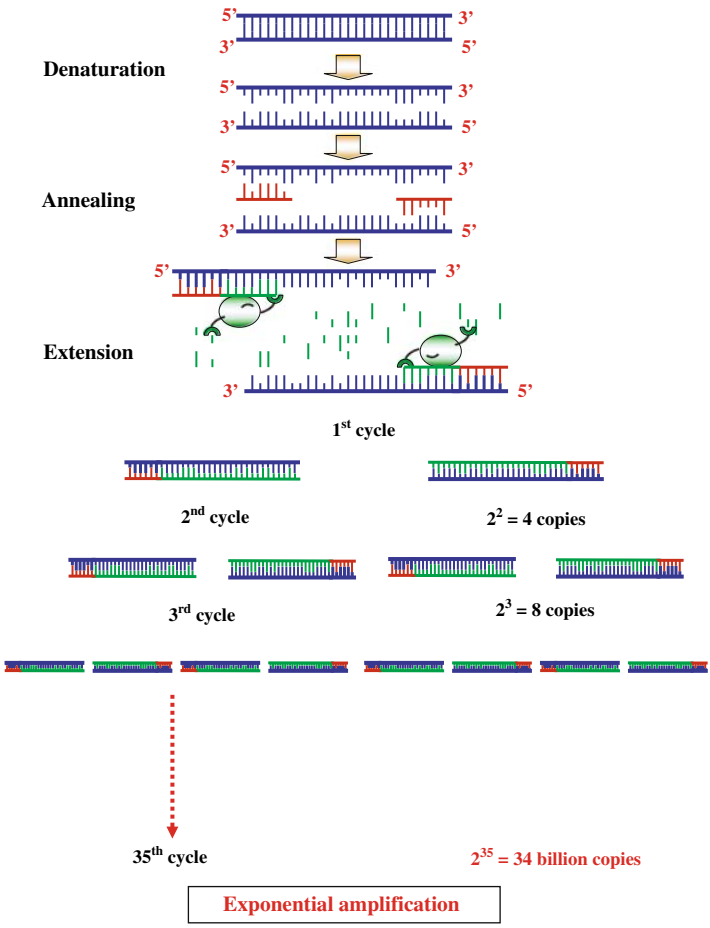


Figure 2. Schematic illustration of exponential amplification. First step: denaturation at 94–96°C. Second step: annealing at (*e.g.*) 65°C. Third step: elongation at 72°C. Fourth step: the first cycle is complete. The two resulting DNA strands make up the template DNA for the next cycle, thus doubling the amount of DNA duplicated for each new cycle

of DNA polymerases. Primer design is extremely important for effective amplification. The primers for the reaction must be very specific for the template to be amplified. Cross-reactivity with non-target DNA sequences results in non-specific amplification. The Tm within the primers should increase from the 5' to the 3' end, in other words, the free energy variation (ΔF) at the 3' end should be maximized using Oligo programs. This makes the priming and the DNA polymerisation by the DNA polymerases more efficient. Additionally, primers must not contain hairpins nor self-anneal. Template length also influences amplification, being particularly efficient for templates in the size range 300 bp to 1 kb. Beyond the upper limit phenomena occur which prevent the PCR products doubling with each cycle, including premature interruptions due to the formation of particular secondary structures, the hybridisation of the neo-synthesized fragments between themselves, *etc*. Nevertheless, it is possible to amplify DNA fragments up to 5 kb long using special polymerases (Fromenty *et al*., 2000). In this case, it is important to take into consideration the half-life of the enzyme which varies considerably between the different DNA polymerases. Half-lives are forty minutes at 100°C for the Taq polymerase, and 23 hours at 95°C and 8 hours at 100°C for the DeepVent polymerase (Jannasch *et al.*, 1992). Above a certain number of cycles (usually 30 to 50 cycles), the rate of amplification gradually decreases. This is due to dNTP depletion, dNTP degradation, primer depletion, the increase in the concentration of PCR products and the inactivation of the polymerase.

5.1. PCR Major Problems

Of the various problems encountered using PCR, some are particularly common:

Contamination is a major problem and is difficult to control (Victor *et al*., 1993). In the years immediately following the advent of PCR considerable occurrence of false positives due to contamination was recorded, an issue of major importance for forensic laboratories. The most frequent source of contamination is previously amplified DNA. The simple act of opening a tube carrying a PCR product (amplicon) can result in the dispersion of an aerosol which contaminates materials destined for use in subsequent amplifications. Contamination can also arise from the presence of samples and/or inadequate laboratory procedures. Thus, it is important to follow good laboratory practice and manipulate PCR products with care (Victor *et al.*, 1993). Spatial organization using two rooms (one for the pre-PCR activities and the other for the post-PCR) can considerably reduce contamination by carryover Avoiding the manipulation of too many DNA samples, the use of pipettes with aerosol barriers, the use of a control (PCR mix without DNA), and the utilization of some commercial devices such as UV light or Uracil-DNA Glycosylase can all diminish DNA contamination. Regarding the latter, PCR can be carried out in the presence of dUTP instead of dTTP, and contamination of the product in subsequent PCR reactions is eliminated by treating the template with uracil DNA glycosylase.

The fidelity of DNA polymerases is a key determinant for accurate amplification (Joyce and Benkovic, 2004). In general, the error rate of DNA polymerases is around one per 10^6 bp (despite their proofreading activity) (Showalter and Tsai, 2002). Variations in this error rate depend on the structure of the DNA molecules and their sequence. Repeated sequences are the more frequent source of errors during amplification. Since Taq DNA polymerase lacks a 3' to 5' exonuclease activity it is unable to proofread and correct the erroneous insertion of a base. This results in a high error rate of approximately 1 in 10,000 bases. However, the error rate of the new generation of DNA polymerases is decreasing.

Non-specific amplifications. Primers used for PCR are usually short, 18 to 25 bases frequently being adequate. The possibility that they hybridise elsewhere rather than to their target sequence is not negligible. Under experimental conditions, where mismatches are common, it has been estimated that 1% of the primers hybridise to their specific sequence whilst the remaining 99% hybridise elsewhere (Kaplan and Delpech, 1994). The resulting non-specific amplifications are not usually detected because they either do not recur or they lead to unidirectional amplification. Gradually increasing the annealing temperature can often resolve problems associated with non-specific binding of primers. The nested PCR technique (see PCR types) is another way to overcome non-specific amplification.

5.2. Practical Modifications to the PCR Technique

Different types of PCR have been developed for different purposes.

Nested PCR is a two PCR run technique (Massung *et al*, 1998). The PCR product of the first run is re-amplified using primers located internally to those used for the first run. This technique is used to avoid contamination due to non-specific amplification and to test primer specificity.

Touchdown PCR is a method that results in increasing specificity of the PCR reaction with increasing cycle numbers (Don *et al*, 1991). The initial annealing temperature should be several degrees above the estimated Tm of the primers. The annealing temperature is then gradually decreased (*e.g.* 1–2°C every second cycle) until it reaches the calculated annealing temperature of the primers or a few degrees lower. Amplification is then continued using this lower annealing temperature.

Inverse PCR (IPCR), described by Ochman *et al.* (1988), is a method for the rapid *in vitro* amplification of DNA sequences that flank a region of known sequence. In this method the primers are placed in the opposite orientation to that usually used, and the template is a restriction fragment that has been self-ligated to form a circle. Inverse PCR has many applications in molecular genetics, including the amplification and identification of sequences flanking transposable elements.

Quantitative PCR (Q-PCR) or **Real-time PCR** is used to rapidly measure the quantity of a PCR product thus providing an indirect but sensitive method for quantifying the starting amounts of template (Lovatt, 2002). The signal from

a fluorescent reporter increases in direct proportion to the amount of PCR product generated as the reaction proceeds. Some of its applications include the quantification of gene expression, drug therapy efficacy and drug monitoring, DNA damage measurement, viral quantification, mitochondrial DNA studies and methylation detection (Kaltenboeck and Wang, 2005).

Amplified fragment-length polymorphism (AFLP) or its fluorescent version (fAFLP) is a PCR-based fingerprinting technology (Shenggi *et al*., 2002). In its most basic form AFLP involves the restriction endonuclease cutting of genomic DNA followed by ligation of adaptors complementary to the restriction sites and selective PCR amplification of a subset of the adapted restriction fragments. These fragments are size-fractionated and visualized on denaturing polyacrylamide gels using autoradiographic or fluorescence methodologies.

Degenerate PCR The major difference in this technique to that of normal PCR is the use of mixed primers to amplify an unknown gene sequence. Degenerate PCR has proven to be a very powerful tool for finding 'new' genes or gene families (Wang *et al*., 2003). Most genes come in families that share structural similarities. By aligning the sequences from a number of related proteins, conserved and variable regions can be identified. Based on this information sequences encoding conserved protein motifs can be used to design degenerate PCR primers.

Thermal asymmetric interlaced (TAIL-) PCR is used for the recovery of DNA fragments adjacent to known sequences. It utilizes a set of nested sequence-specific primers together with a shorter arbitrary degenerate primer. TAIL-PCR is an efficient technique for amplifying insert ends from yeast artificial chromosome (YAC) and P1 clones (Singer and Burke, 2003). Highly specific amplification is achieved without resorting to complex manipulations before or after PCR.

Meta-PCR is a simple, versatile, and powerful method for generating chimeric DNA molecules. Up to five PCR amplifiable fragments can be combined to form a single linear amplimer. The Meta-PCR reaction is self-assembling and takes place in two coupled stages carried out in a single reaction vessel. The order of fragments is reproducible and determined by primer design. Meta-PCR is likely to be useful to clinical molecular diagnostic laboratories, helping them to fulfil demand for rapid and accurate screening for point mutations in large multi-exon genes (Wallace *et al.*, 1999).

Multiplex PCR is a screening strategy that enables the simultaneous amplification of multiple DNA targets by using several pairs of primers. Many factors can impair this technique. Most importantly all primer pairs must have the same effective Tm. In addition, the amount of enzyme required is greater than that for uniplex PCR. For example, a Taq DNA polymerase concentration (with an appropriate increase in $MgCl_2$ concentration) four to five times greater than required in uniplex PCR is necessary to achieve optimal amplification. Multiplex PCR is used for the detection of gene mutations and deletions, polymorphism screening, and the identification of virus, bacteria, fungi and parasites (Elnifro *et al*., 2000) where repeated PCR amplification of the same targets is required.

Ligation-mediated PCR (LMPCR) technology can be divided into nine steps (Drouin *et al*., 2001) (Fig. 3): (I) conversion of modified bases to single-stranded breaks (SSBs); (II) heat denaturation of genomic DNA; (III) hybridisation and extension of a gene-specific oligonucleotide (primer 1) for either DNA strand to produce DNA molecules with an unknown double-stranded blunt 3' end; (IV) ligation of an asymmetrical double-stranded DNA linker to provide a common known sequence; (V–VI) linear and exponential PCR amplifications using a gene-specific nested oligonucleotide (primer 2) and the linker-specific oligonucleotide (linker primer); (VII) size-fractionation of the PCR products on a polyacrylamide sequencing gel and transfer of the DNA to a nylon membrane by electroblotting; (VIII) hybridisation with a gene-specific labelled probe generated using primer 2 or a nested oligonucleotide with a PCR product corresponding to the sequence to be analysed; and (IX) washing of the membrane and revealing of the sequence ladder by autoradiography (Drouin *et al*., 2001).

LMPCR technology is an extremely sensitive and specific genomic sequencing technique which has been successfully applied for over a decade by many groups to *in vivo* DNA–protein interaction analysis (footprinting), DNA damage mapping, methylation analysis and nucleosome positioning (Drouin *et al*., 2001).

5.3. Applications of PCR

PCR technology allows the utilisation of very small specimens of genetic material, even from just one cell, copying it repeatedly to generate a test sample sufficiently amplified to be able to detect the presence or absence of a viral, bacterial or other specific DNA sequence. It is indispensable for the amplification of DNA samples when there is not enough material available to analyse by other methods (*e.g.* DNA from crime scenes, archaeological samples, *etc*).

Medical diagnosis: PCR can detect and identify bacteria and viruses that cause infections such as tuberculosis, chlamydia, viral meningitis, viral hepatitis, HIV, cytomegalovirus and many others. Once primers are designed for a specific sequence, PCR can detect the presence or absence of the corresponding pathogen in a patient's blood or tissues (McIver *et al*., 2005).

Genetic testing: PCR is also used to determine whether patients carry a genetic mutation that could be inherited by their offspring (*e.g.* the mutation that causes cystic fibrosis (Ivan den Bergh and Martens, 2003)) or present a disease risk in patients themselves (*e.g.* a mutation in the BRCA1 gene that predisposes a woman to breast or ovarian cancer (Angioli *et al*., 1998)). Parts of the gene are amplified by PCR and then sequenced to look for mutations.

Genome sequencing: Using random primers the entire genome of an organism can be amplified in pieces. Once amplified the pieces must be sequenced, and then the overlapping sequences can be joined by computer algorithms to determine the genome sequence (Csakó, 2006).

Mutagenesis: Mutations can be introduced into copied DNA sequences in two fundamentally different ways during the PCR process. Site-directed mutagenesis

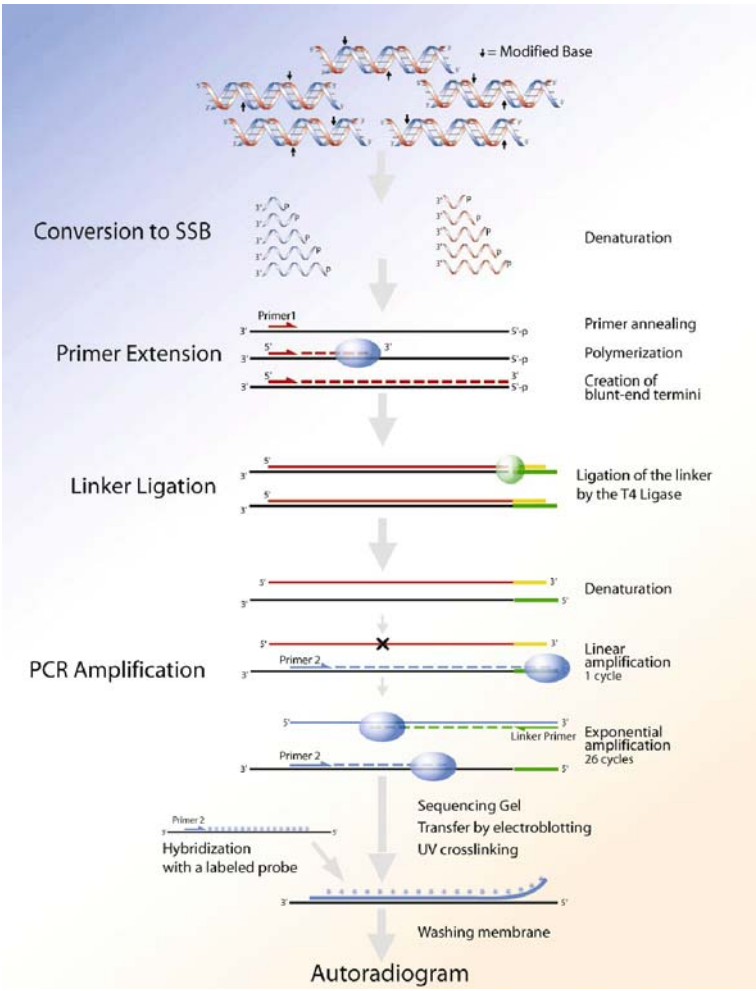


Figure 3. Overview of the different steps in the LMPCR protocol. Step I: specific conversion of modified bases to phosphorylated single-strand breaks; Step II: denaturation of genomic DNA; Step III: annealing and extension of primer 1 (although both strands can be studied, each LMPCR protocol only involves the analysis of either the non-transcribed strand or the transcribed strand); Step IV: ligation of the linker; Step V: first cycle of PCR amplification, this cycle is a linear amplification because only the gene-specific primer 2 can anneal; Step VI: cycle 2 to 22 of exponential PCR amplification of gene-specific fragments with primer 2 and the linker primer (the longer oligonucleotide of the linker); Step VII: separation of the DNA fragments on a sequencing gel, transfer of the sequence ladder to a nylon membrane by electroblotting; Step VIII: preparation of single-stranded probe and hybridisation; Step IX: washing of membrane and visualization of the sequence ladder

allows the experimenter to introduce a mutation at a specific location on the DNA strand (Barettino *et al*., 1994). The desired mutation is usually incorporated in the primers used for the PCR program. By contrast, random mutagenesis is based on the use of error-prone polymerases in the PCR process (Biles and Connolly, 2004). In the case of random mutagenesis, the location and nature of the mutations cannot be controlled. One application of random mutagenesis is to analyse structure-function relationships in proteins.

Analysis of ancient DNA: Using PCR it becomes possible to analyse DNA that is thousands of years old. PCR techniques have been successfully used on animals, such as a forty-thousand-year-old mammoth, and on human DNA (Dissing *et al*., 2006), in applications ranging from the analysis of Egyptian mummies (Chastel, 2004) to the identification of a Russian Tsar.

5.4. Choosing a DNA Polymerase

The choice of a DNA polymerase is determined mainly by the goal of a given experiment (see Table 2). Clearly defining the objective of an experiment considerably helps in the judicious choice of a DNA polymerase.

The size of the targeted sequence: some polymerases are good at quickly and efficiently amplifying short sequences(<1 kb) but cannot amplify longer ones. For example, Taq and Pfu are good for short sequences while rTth DNA polymerase XL and phi29 DNA polymerase are good for long ones, up to 40 kb. One may wish to use a DNA polymerase lacking terminal transferase activity to exactly amplify the size of a sequence.

Duration of the polymerisation step in the PCR cycle: since different polymerases have different rates of polymerisation, the duration of the polymerisation step must take into account both the rate of polymerisation and the length of the sequence to be amplified. Unfortunately, the rates of polymerisation of several DNA polymerases are not available in the literature nor provided by the companies that produce them.

The starting amount of template: certain polymerases require fewer template DNA molecules than others (Angers *et al*., 2001; Vigneault and Drouin, 2005).

Maximum PCR product quantity: the amount of the desired product increases exponentially with cycle number (exponential amplification phase) but this rate decreases at high cycle numbers. This decrease is often related with the half-life of the polymerase used (Dietrich *et al*., 2002) (Table 2) or to dNTP depletion (Markoulatos *et al*., 2002). In addition, the quality of the amplified product is often compromised at cycle numbers beyond the exponential phase (Martell *et al*., 1999). Usually, 25 to 30 cycles is sufficient to get 1 μg of PCR product.

The quality of the template: impure DNA can compromise the quality of the amplified product. Where possible it is also useful to know the sequence of the template. The presence of numerous GC rich sequences and particular secondary structure strongly influences DNA polymerase activity (Hube *et al*., 2005). The use of DMSO

(DiMethylSulfOxide), TMSO (TriMethylSulfOxide) and betaine assists polymerases to progress through these GC-rich sequences (Chakrabarti and Schutt, 2002).

The quality of the PCR product: a DNA polymerase with a very low error rate is recommended for sequencing (Table 2) (Kunkel and Bebenek, 2000). Some DNA polymerases are very useful for cloning because of their terminal transferase activity.

Financial aspects: primarily related to the cost of DNA polymerase. Table 2 gives choices of polymerase to use and what compromises may be made to lower costs.

Table 3 summarizes suggestions for DNA polymerases to use based on the limitations/conditions associated with the aim of a given PCR. Descriptions of the different DNA polymerases are reported in section 4. It is very difficult to index the weaknesses of the different DNA polymerases since this kind of information is not generally published.

5.5. A Practical Example: Choice of DNA Polymerases for LMPCR

LMPCR (Drouin *et al*, 2001) involves the PCR amplification of a mixture of genomic DNA fragments of different sizes. During the LMPCR procedure, DNA polymerases are required for two steps: primer extension (PE) and PCR amplification. For the PE step, the best DNA polymerase to use would be one that (1) is thermostable and very efficient, (2) has no terminal transferase activity, (3) is able to efficiently polymerise about 0.75 kb of DNA even when the DNA is very GC rich, and (4) is able to polymerise through any particular DNA secondary structures. For the PCR step, the ideal polymerase would be (1) thermostable, (2) very efficient,

Table 3. Suggested uses for various DNA polymerases

Options	Suggested DNA polymerases
Starting with a very low template concentration	phi29 DNA polymerase BioTherm^{tm}DNA polymerase
Amplify a very long DNA fragment	5 to 40 kb: rThh DNA polymerase up to 70 kb: phi29 DNA polymerase
Amplify a GC-rich sequence with a particular secondary structure	Herculase® Enhanced rBst DNA Polymerase
Fidelity	Isis proofreading DNA polymerase
Using non-optimised primers	SurePRIME^{tm}DNA polymerase
Rapid amplification	SpeedSTAR^{tm} HS DNA polymerase
Cloning	Phusion^{TM} DNA polymerase
Amplify for a long time	Deep Vent®DNA polymerase
Amplify the exact size of the DNA fragment	pfu DNA polymerase
3'-OH end labelling	Taq DNA polymerase
TA cloning	Taq DNA polymerase
Vector construction from single DNA strands	DNA polymerase I

(3) able to amplify indiscriminately a mixture of DNA fragments of different lengths (between 50 and 750 bp) and of varying GC-richness (from 5 to 95%), and (4) able to efficiently resolve particular DNA secondary structures. We found cloned Pfu DNA polymerase that corresponds to Pfu exo$^-$ to be the best enzyme for the PE and PCR steps of LMPCR (Angers *et al.*, 2001).

6. FUTURE PROSPECTS AND CONCLUSIONS

Alternative methods are beginning to emerge from the 20-year history of PCR amplification techniques. Whilst thermocycling is an essential component of all PCR protocols - required for DNA denaturation and primer annealing - it makes the methods technically more complex, necessitating thermostable enzymes that can survive not only the denaturation steps but also 'hot start' procedures designed to avoid non-specific amplifications. Isothermal methods could simplify amplification, reducing both the time required and the cost of the process. Currently, access to such methods is limited but developments are progressing in this area (Csakd, 2006). Helicase-dependent amplification (HDA) is a new isothermal DNA amplification method in which a single temperature (37°C) is used for the whole process. HDA uses a DNA helicase to generate single-stranded templates to allow annealing of the primer; subsequent elongation is done by a DNA polymerase. *Thermoanaerobacter tengcongensis* is an anaerobic eubacterium from which a mutant helicase has been isolated and cloned that is stable and active from 45°C to 65°C. This property has permitted the development of a thermophilic helicase-dependent amplification (tHDA) assay. Its principle is to selectively amplify target sequences between 60°C and 65°C. This assay is more efficient than HDA because it does not need the participation of accessory proteins and shows increased amplification sensitivity (Csakd, 2006).

New technologies (*e.g.* microfluidics, micro-thermal and micro-propulsion technologies, lab-on-a-chip technology) coupled with developments in biochemistry and computer sciences provide new ways for nucleic acid based analyses and testing. The PCRJet (Megabase Research Products, Lincoln, NE) is an excellent example to illustrate the important advances in amplification methods. PCRJet can generate 100 to 600 bp long amplicons after 30 cycles over the course of just 2 minutes; to amplify 2331 bp in 30 cycles takes about 10 minutes. This impressive performance makes PCRjet the fastest instrument ever developed for DNA amplification. PCRjet uses the very fast KOD DNA polymerase (from *Thermococcus kodakaraensis*) which in one second can synthesize 300 bp (Moore, 2005).

Molecular biology methods are becoming extremely sensitive and specific, and their increasing availability and speed will be important not only for laboratories specialised in disease diagnosis but also in basic research and biotechnology-based industries. Molecular diagnoses using powerful and easy methodologies will become routine in the prevention and treatment of numerous types of genetic, cancerous and infectious diseases. Full automation of these methods will provide rapid and accurate tests from minimal amounts of specimen at the lowest possible cost.

ACKNOWLEDGMENTS

The authors are grateful to Mrs. Nathalie Bastien, Isabelle Paradis and Mr. Macoura Gadji for their help with the writing of this chapter, and Dr. Gerald P. Holmquist for carefully editing the manuscript. The work done in the Laboratory of Dr. Régen Drouin was supported by a grant from the Canadian Genetic diseases Network (MRC/NSERC NCE Program). Dr. Régen Drouin holds the Canada Research Chair in 'Genetics, Mutagenesis, and Cancer'.

REFERENCES

Angers, M., Cloutier, J.F., Castonguay, A. and Drouin, R. (2001) Optimal conditions to use Pfu exo(-) DNA polymerase for highly efficient ligation-mediated polymerase chain reaction protocols. Nucleic Acids Res. *29*, E83.

Angioli, R., Estape, R., Mason, M. and Penalver, M. (1998) Hereditary and sporadic ovarian cancer: genetic testing and clinical implications (review). Int. J. Oncol. *12*, 1029–1034.

Barettino, D., Feigenbutz, M., Valcarcel, R. and Stunnenberg, H.G. (1994) Improved method for PCR-mediated site-directed mutagenesis. Nucleic Acids Res. *22*, 541–542.

Beard, W.A. and Wilson, S.H. (2006) Structure and mechanism of DNA polymerase Beta. Chem. Rev. *106*, 361–382.

Bebenek, K. and Kunkel, T.A. (2004) Functions of DNA polymerases. Adv. Protein Chem. *69*, 137–165.

Biles, B.D. and Connolly, B.A. (2004) Low-fidelity *Pyrococcus furiosus* DNA polymerase mutants useful in error-prone PCR. Nucleic Acids Res. *32*, e176.

Blanco, L., Bernad, A., Lazaro, J.M., Martin, G., Garmendia, C. and Salas, M. (1989) Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. J. Biol. Chem. *264*, 8935–8940.

Bohlke, K., Pisani, F.M., Rossi, M. and Antranikian, G. (2002) Archaeal DNA replication: spotlight on a rapidly moving field. Extremophiles *6*, 1–14.

Braithwaite, D.K. and Ito, J. (1993) Compilation, alignment, and phylogenetic relationships of DNA polymerases. Nucleic Acids Res. *21*, 787–802.

Brock, T.D. and Freeze, H. (1969) *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. J. Bacteriol. *98*, 289–297.

Burgers, P.M., Koonin, E.V., Bruford, E., Blanco, L., Burtis, K.C., Christman, M.F., Copeland, W.C., Friedberg, E.C., Hanaoka, F., Hinkle, D.C., Lawrence, C.W., Nakanishi, M., Ohmori, H., Prakash, L., Prakash, S., Reynaud, C.A., Sugino, A., Todo, T., Wang, Z., Weill, J.C. and Woodgate, R. (2001) Eukaryotic DNA polymerases: proposal for a revised nomenclature. J. Biol. Chem. *276*, 43487–43490.

Camps, M. and Loeb, L.A. (2004) When pol I goes into high gear: processive DNA synthesis by pol I in the cell. Cell Cycle *3*, 116–118.

Chakrabarti, R. and Schutt, C.E. (2002) Novel sulfoxides facilitate GC-rich template amplification. Biotechniques *32*, 866, 868, 870–872, 874.

Chang, L.M. and Bollum, F.J. (1986) Molecular biology of terminal transferase. CRC Crit. Rev. Biochem. *21*, 27–52.

Chastel, C. (2004) When the Egyptian mummies are speaking about the infections that have made them ill. Hist. Sci. Med. *38*, 147–155.

Copeland, W.C. and Longley, M.J. (2003) DNA polymerase gamma in mitochondrial DNA replication and repair. ScientificWorldJournal *3*, 34–44.

Csako, G. (2006) Present and future of rapid and/or high-throughput methods for nucleic acid testing. Clin. Chim. Acta *363*, 6–31.

D'Urso, G. and Nurse, P. (1997) *Schizosaccharomyces pombe* cdc20+ encodes DNA polymerase epsilon and is required for chromosomal replication but not for the S phase checkpoint. Proc. Natl. Acad. Sci. USA *94*, 12491–12496.

Dietrich, J., Schmitt, P., Zieger, M., Preve, B., Rolland, J.L., Chaabihi, H. and Gueguen, Y. (2002) PCR performance of the highly thermostable proof-reading B-type DNA polymerase from *Pyrococcus abyssi.* FEMS Microbiol. Lett. *217*, 89–94.

Dissing, J., Binladen, J., Hansen, A., Sejrsen, B., Willerslev, E. and Lynnerup, N. (2006) The last Viking King: A royal maternity case solved by ancient DNA analysis. Forensic Sci. Int. In press.

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K. and Mattick, J.S. (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. *19*, 4008.

Drouin, R., Therrien, J.P., Angers, M. and Ouellet, S. (2001) *In vivo* DNA analysis. Methods Mol. Biol. *148*, 175–219.

Elnifro, E.M., Ashshi, A.M., Cooper, R.J. and Klapper, P.E. (2000) Multiplex PCR: optimization and application in diagnostic virology. Clin. Microbiol. Rev. *13*, 559–570.

Franklin, M.C., Wang, J. and Steitz, T.A. (2001) Structure of the replicating complex of a pol alpha family DNA polymerase. Cell *105*, 657–667.

Fromenty, B., Demeilliers, C., Mansouri, A. and Pessayre, D. (2000) *Escherichia coli* exonuclease III enhances long PCR amplification of damaged DNA templates. Nucleic Acids Res. *28*, E50.

Graves, S.W., Johnson, A.A. and Johnson, K.A. (1998) Expression, purification, and initial kinetic characterization of the large subunit of the human mitochondrial DNA polymerase. Biochemistry *37*, 6050–6058.

Gueguen, Y., Rolland, J.L., Lecompte, O., Azam, P., Le Romancer, G., Flament, D., Raffin, J.P. and Dietrich, J. (2001) Characterization of two DNA polymerases from the hyperthermophilic euryarchaeon *Pyrococcus abyssi.* Eur. J. Biochem. *268*, 5961–5969.

Hube, F., Reverdiau, P., Iochmann, S. and Gruel, Y. (2005) Improved PCR method for amplification of GC-rich DNA sequences. Mol. Biotechnol. *31*, 81–84.

Hubscher, U., Maga, G. and Spadari, S. (2002) Eukaryotic DNA polymerases. Annu. Rev. Biochem. *71*, 133–163.

Jannasch, H.W., Wirsen, C.O., Molyneaux, S.J. and Langworthy, T.A. (1992) Comparative Physiological Studies on Hyperthermophilic Archaea Isolated from Deep-Sea Hot Vents with Emphasis on Pyrococcus Strain GB-D. Appl. Environ. Microbiol. *58*, 3472–3481.

Joyce, C.M. and Benkovic, S.J. (2004) DNA polymerase fidelity: kinetics, structure, and checkpoints. Biochemistry *43*, 14317–14324.

Kaltenboeck, B. and Wang, C. (2005) Advances in real-time PCR: application to clinical laboratory diagnostics. Adv. Clin. Chem. *40*, 219–259.

Kannouche, P. and Stary, A. (2003) Xeroderma pigmentosum variant and error-prone DNA polymerases. Biochimie *85*, 1123–1132.

Kaplan, M.C. and Delpech, M. (1994) Les techniques d'amplification élective *in vitro* (PCR, etc.). In Médecine-Sciences (eds) *Biologie moléculaire et médecine*, Flammarion, Paris, *2*, 558–572.

Kaplan, M.C. and Delpech, M. (1994) Les outils enzymatiques du génie génétique. In Médecine-Sciences (eds) *Biologie moléculaire et médecine*, Flammarion, Paris, *2*, 573–584.

Kornberg, A. (1957) Enzymatic synthesis of deoxyribonucleic acid. Harvey Lect. *53*, 83–112.

Kunkel, T.A. and Bebenek, K. (2000) DNA replication fidelity. Annu. Rev. Biochem. *69*, 497–529.

Lewin, B. (2004) DNA replication. In (eds) Genes VIII, Pearson Prentice Hall, Upper Saddle River, NJ, *8*, 387–417.

Loeb, L.A. (2001) A mutator phenotype in cancer. Cancer Res. *61*, 3230–3239.

Lovatt, A. (2002) Applications of quantitative PCR in the biosafety and genetic stability assessment of biotechnology products. J. Biotechnol. *82*, 279–300.

Markoulatos, P., Siafakas, N. and Moncany, M. (2002) Multiplex polymerase chain reaction: a practical approach. J. Clin. Lab. Anal. *16*, 47–51.

Martell, M., Gomez, J., Esteban, J.I., Sauleda, S., Quer, J., Cabot, B., Esteban, R. and Guardia, J. (1999) High-throughput real-time reverse transcription-PCR quantitation of hepatitis C virus RNA. J. Clin. Microbiol. *37*, 327–332.

Massung, R.F., Slater, K., Owens, J.H., Nicholson, W.L., Mather, T.N., Solberg, V.B. and Olson, J.G. (1998) Nested PCR assay for detection of granulocytic ehrlichiae. J. Clin. Microbiol. *36*, 1090–1095.

McIver, C.J., Jacques, C.F., Chow, S.S., Munro, S.C., Scott, G.M., Roberts, J.A., Craig, M.E. and Rawlinson, W.D. (2005) Development of multiplex PCRs for detection of common viral pathogens and agents of congenital infections. J. Clin. Microbiol. *43*, 5102–5110.

Moore, P. (2005) PCR: replicating success. Nature *435*, 235–238.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. (1986) Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. Cold Spring Harb. Symp. Quant. Biol. *51*(1), 263–273.

Ochmanm, H., Gerber, A. S. and Hartl, D. L. (1988) Genetic applications of an inverse polymerase chain reaction. Genetics *120*, 621–623.

Ponamarev, M.V., Longley, M.J., Nguyen, D., Kunkel, T.A. and Copeland, W.C. (2002) Active site mutation in DNA polymerase gamma associated with progressive external ophthalmoplegia causes error-prone DNA synthesis. J. Biol. Chem. *277*, 15225–15228.

Pospiech, H., Kursula, I., Abdel-Aziz, W., Malkas, L., Uitto, L., Kastelli, M., Vihinen-Ranta, M., Eskelinen, S. and Syvaoja, J.E. (1999) A neutralizing antibody against human DNA polymerase epsilon inhibits cellular but not SV40 DNA replication. Nucleic Acids Res. *27*, 3799–3804.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Sawaya, M.R., Pelletier, H., Kumar, A., Wilson, S.H. and Kraut, J. (1994) Crystal structure of rat DNA polymerase beta: evidence for a common polymerase mechanism. Science *264*, 1930–1935.

Shengqi, W., Xiaohong, W., Suhong, C. and Wei, G. (2002) A new fluorescent quantitative polymerase chain reaction technique. Anal. Biochem. *309*, 206–211.

Showalter, A.K. and Tsai, M.D. (2002) A reexamination of the nucleotide incorporation fidelity of DNA polymerases. Biochemistry *41*, 10571–10576.

Singer, T. and Burke, E. (2003) High-throughput TAIL-PCR as a tool to identify DNA flanking insertions. Methods Mol. Biol. *236*, 241–272.

Steitz, T.A. (1998) A mechanism for all polymerases. Nature *391*, 231–232.

van den Bergh, F.A. and Martens, A. (2003) Diagnosis of cystic fibrosis; simple genotyping to rule out the disease preferable to starting with the sweat test. Ned. Tijdschr. Geneeskd. *147*, 1001–1005.

Vanderstraeten, S., Van den Brule, S., Hu, J. and Foury, F. (1998) The role of 3'-5' exonucleolytic proofreading and mismatch repair in yeast mitochondrial DNA error avoidance. J. Biol. Chem. *273*, 23690–23697.

Victor, T., Jordaan, A., du Toit, R. and Van Helden, P.D. (1993) Laboratory experience and guidelines for avoiding false positive polymerase chain reaction results. Eur. J. Clin. Chem. Clin. Biochem. *31*, 531–535.

Vigneault, F. and Drouin, R. (2005) Optimal conditions and specific characteristics of Vent exo- DNA polymerase in ligation-mediated polymerase chain reaction protocols. Biochem. Cell. Biol. *83*, 147–165.

Wallace, A.J., Wu, C.L. and Elles, R.G. (1999) Meta-PCR: a novel method for creating chimeric DNA molecules and increasing the productivity of mutation scanning techniques. Genet. Test *3*, 173–183.

Wang, H.Z., Zhou, X.F., Song, Z.X., Liu, W.G. and Zeng, X.L. (2003) Degenerate PCR and its application in gene cloning. Yi Chuan *25*, 201–204.

Ye, S.Y. and Hong, G.F. (1987) Heat-stable DNA polymerase I large fragment resolves hairpin structure in DNA sequencing. Sci. Sin. [B] *30*, 503–506.

Zhou, M.Y. and Gomez-Sanchez, C.E. (2000) Universal TA cloning. Curr. Issues Mol. Biol. *2*, 1–7.

CHAPTER 23

PROKARYOTIC REVERSE TRANSCRIPTASES

BERT C. LAMPSON*

Department of Health Sciences, East Tennessee State University, Johnson City, TN, USA
**lampson@mail.etsu.edu*

1. INTRODUCTION

The reverse transcription of the genetic code of an RNA molecule back into a complementary DNA copy (cDNA) is usually accomplished by a specially dedicated RNA-dependent DNA polymerase called a reverse transcriptase (RT). Reverse transcription is presumed to have been a central event in the transition from the theorized ancient "RNA world" of life, in which RNA molecules served as both the source of genetic information, as well as, catalytic functions for cellular life, to the present day DNA-RNA-protein world. Even today reverse transcription continues to occur in most organisms from human cells to bacteria. For example, a wide assortment of genetic elements found in plant, animal, and microbial cells use reverse transcription for at least part of their replication or mobility. These include RNA viruses, DNA viruses, transposons, introns, and mitochondrial plasmids (Eickbush, 1994; Eickbush and Malik, 2002). More astonishing than the wide variety of retroelements that encode a RT is the colossal number of repetitive retrosequences, such as the Alu sequences, that have been generated by these elements in many eukaryotic genomes. These retrosequences are DNAs that usually do not code for RT but have clearly been produced by reverse transcription of an RNA molecule (Kazazian, 2004). In addition to the production of these "parasitic" DNAs, reverse transcription also serves a vital function for most eukaryotic organisms. Here the essential enzyme telomerase, which functions to maintain the telomere ends of chromosomes, is a type of RT (Lue, 2004).

Most RTs are placed in two very broad categories based on the phylogenetic analysis of their amino acid sequence and the type of retroelement that codes for these polymerases. The first group is found in eukaryotic cells and is usually called LTR-containing retroelements because their DNA is flanked by *l*ong *t*erminal *r*epeat

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 403–420

sequences. These elements include retroviruses and the virus-like retrotransposons called Ty in yeast cells. The second group is called non-LTR retroelements because they are not flanked by long terminal repeats and are a very diverse collection of elements found in both eukaryotic and prokaryotic organisms. These include the mobile group II introns, the L1 retrotransposons, and the various retroelements found in bacteria. Telomerase is phylogenetically related to the RTs from this group and is thus also included in this second category (Eickbush and Malik, 2002).

Among the best characterized RTs are the polymerases from retroviruses and include the detailed crystal structure of the RT from the HIV-1 virus (Steitz, 1998). The focus of this chapter is the recently discovered RTs found in bacteria and the interesting genetic elements that encode them. These prokaryotic RTs are considered to be the ancestors of all retroelements found in eukaryotic genomes based on phylogenetic comparisons (Toor *et al*, 2001). In addition, the bacterial RTs are a very diverse group of proteins with a number of novel properties. Thus, these bacterial RTs represent an emerging new resource for many potential nucleic acid based technologies like RT-PCR.

Prokaryotic RTs generally fall into one of three different types depending on the type of retroelement DNA that codes for these proteins. These groups or types are also in agreement with phylogenetic groupings for RTs determined by comparing their amino acid sequences. The three types of retroelements are (1) the group II introns found in both eubacterial and archaeal genomes, (2) retrons, which are also found in eubacteria and some archaea, and (3) the diversity generating retroelements that have thus far been found in the eubacteria.

2. GROUP II INTRONS

Among the best understood retroelements of prokaryotic organisms are the group II introns. The group II introns were initially discovered in the genomes of organelles like chloroplasts and the mitochondria of yeast (Michel and Lang, 1985). Using degenerate primers and a PCR screening method, group II introns were subsequently discovered in a variety of different bacteria as well (Ferat and Michel, 1993). Like any intron DNA, group II introns interrupt a gene (although this occurs only occasionally in bacteria) and must be spliced out of the corresponding RNA transcript to yield a mature RNA encoding a functional gene. In the initial RNA transcript, independent of any host cell splicing factor, group II introns form an autocatalytic RNA that can self-splice from the RNA molecule. The excised intron RNA forms a characteristic lariat structure by a mechanism similar to the splicing of nuclear introns by the eukaryotic spliceosome (Saldanha *et al*., 1993).

In addition to self-splicing, most group II intron DNAs found in bacteria contain an open reading frame (ORF). The product of this ORF gene is a multifunctional protein with (1) maturase activity that aids in intron splicing (Matsurra *et al*, 2001), (2) RT activity that converts the spliced intron RNA into a cDNA copy and, in some cases (3) a DNA endonuclease that helps the intron insert into a new site in a DNA molecule (Martinez-Abarca and Toro, 2000; Zimmerly *et al*., 2001). This intron

encoded protein allows some intron DNAs to act as retrotransposable elements that can insert into new DNA molecules (Cousineau *et al.*, 1998, 2000).

2.1. Intron RNA

At 2 to 2.5 kilobase pairs in size, group II intron DNA from bacteria codes for two important molecules required for intron mobility and splicing. First is the intron RNA, which will fold into a ribozyme required for self-splicing. And second is the encoded protein, which aids in splicing and is also required for retrotransposition (Fig. 1).

Group II intron RNA must fold into a series of complex secondary stem-loop structures in order to form the autocatalytic ribozyme required for self-splicing. These stem-loop structures are designated domain I through VI (Fig. 1). The catalytic core of the intron ribozyme appears to be centered around stem-loop domain V, which is also the most conserved in its primary nucleotide sequence among the many different intron RNAs (Toor *et al.*, 2001; Lambowitz and Zimmerly, 2004). A small stem-loop structure within domain IV appears to function as a high affinity binding site for the intron RT protein. Here the intron protein binds to its own intron RNA and helps ensure that the intron RNA folds efficiently into a catalytically functional ribozyme (Singh *et al.*, 2002). Finally, domain structure VI contains the so called "bulging A" (large A in Fig. 1) base. This is the internal adenine nucleotide that will form the 2'–5' branch linkage with the 5' terminal base of the spliced intron to form the lariat intermediate RNA.

Although the basic architecture of the intron RNA described above is shared by all group II introns, there is a great deal of structural variability especially among the bacterial introns. There are currently five recognized subgroups or classes of group II introns found in bacteria (bacterial class A–E) based on the high degree of variation in their folded RNA structure (Toor *et al.*, 2001; Lambowitz and Zimmerly, 2004).

2.2. Intron Protein

Unlike the group II introns found in organelle genomes, almost all group II introns found in bacteria have an ORF capable of producing an intron protein, anywhere from about 415 to 600 amino acids in size. This intron protein is a multifunctional protein with several functionally defined regions or domains (Fig. 2). Beginning at the N-terminus and encompassing most of the protein is a region required for RT activity. Seven blocks (domains 1-7) of conserved amino acid sequence define this RT domain and are shared among all other RTs including the retroviral RT of HIV (Xiong and Eickbush, 1990). In addition, this region also defines the so called fingers and palm regions of the crystal structure of HIV RT suggesting that this bacterial RT contains a similar folded structure (Steitz, 1999; Blocker *et al.*, 2005). An additional block of conserved amino acids that is shared among the bacterial intron RTs and other RTs from non-LTR type transposons, is designated 0 (Blocker *et al.*, 2005).

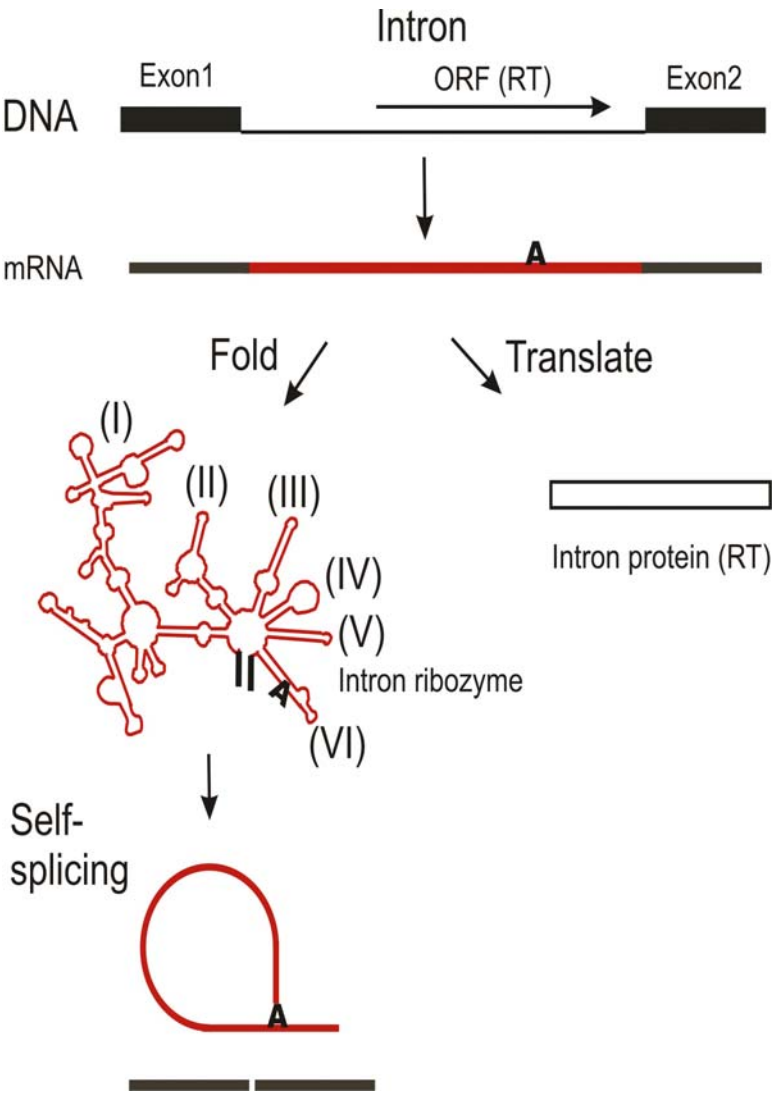


Figure 1. Group II introns in bacteria are mobile DNAs. From a mRNA transcript of the group II intron, two important molecules are derived that are required for mobility. First, an intron ORF is translated to yield a protein with RT activity. Second, the mRNA molecule folds into a complex secondary structure containing six structural domains designated I-IV (in parentheses). This folded RNA molecule (with help from the intron protein) forms an active ribozyme that can self-splice. Splicing joins the exon sequences of the mRNA molecule together and removes the intron sequences as a lariat molecule with the 5'-end of the intron RNA linked to the 2' position of an internal adenine nucleotide (large A). The intron protein binds to the spliced intron RNA to form a ribonucleoprotein complex that will mobilize the element to new target sites

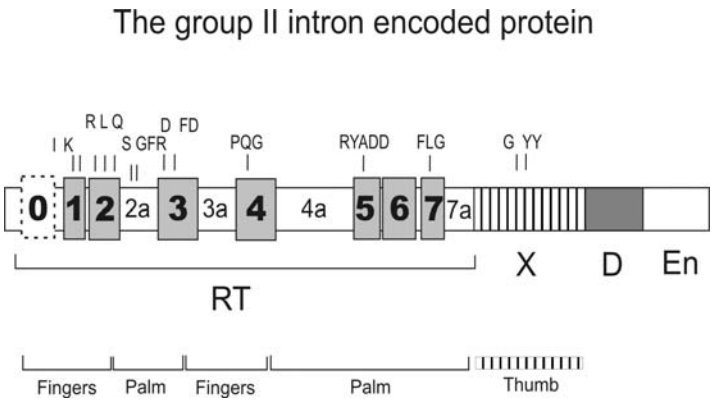


Figure 2. The group II intron protein is a multi-functional enzyme. Functional regions of the intron protein are shown as boxes in the diagram. For example, at the N-terminal half of the protein (left half of the long open rectangle) is the RT domain defined as seven blocks of conserved amino acids (numbered boxes 1-7) shared by all RTs. Some highly conserved amino acids in this domain found among group II intron proteins are shown (single-letter abbreviations above the numbered boxes). Structural modeling and protease cleavage analysis indicates that the RT domain resembles the fingers-palm folded region in the crystal structure of HIV RT (brackets in diagram). Just beyond the RT domain is a functional domain designated "X" that is associated with maturase activity. Evidence indicates that this region folds into a structure resembling the thumb region in the crystal structure of HIV RT. At the C-terminus of the intron are two domains designated D and En. Domain D appears to function in binding DNA and domain En is associated with an endonuclease activity that cleaves DNA

C-terminal to the RT region of the intron protein is a block of amino acids forming a domain called X. This sequence of amino acids is conserved among group II intron proteins but is not found in other RT proteins (Zimmerly *et al.*, 2001). In addition, mutational changes in this region of the protein affect the activity of intron RNA splicing (Mohr *et al.*, 1993). For this reason, domain X is considered to be the region of the protein involved in "maturase" activity. For maturase activity, the intron protein functions as a kind of protein splicing factor. Here, domain X, and parts of the RT region of the intron protein bind to specific regions of the intron RNA and help to fold the intron RNA into a stable ribozyme active site for self-splicing. Although domain X appears to be unique to the group II intron proteins, it nevertheless contains structural features similar to the thumb region in the crystal structure of HIV RT. This includes three α-helices predicted to form in the X domain of group II intron proteins (Blocker *et al.*, 2005), and potentially correspond to the 3 α-helices (α-H, α-I, and α-J) in the thumb of HIV RT.

At the C-terminus of the intron protein are two functional domains designated D and En (Fig. 2). Both of these regions appear to function in different steps that occur during retrotransposition since truncation of the C-terminal end of the protein to remove these domains does not affect splicing of the intron RNA (Guo *et al.*, 1997). The D domain appears to have a DNA binding function that binds to and unwinds DNA at the site of transposition. The terminal En domain resembles the amino acid

sequence of the H-N-H type DNA endonucleases. This is a member of the so called homing endonucleases that promote the mobility of intron DNA, of both group I and group II introns, by recognizing and cleaving DNA at intron homing sites (see intron mobility below). There are several different types of these endonucleases such as the H-N-H family. The H-N-H family of enzymes contain conserved histidine amino acids flanking an asparagine, and include colicins, phage encoded endonucleases and the intron homing enzymes (Belfort and Roberts, 1997). This endonuclease activity carries out the second-strand cleavage event during retro-mobility (Guo *et al.*, 1997). Many of the group II intron proteins from bacteria do not contain the En domain and thus carry out retro-mobility without this endonuclease activity by a slightly different mechanism (Ichiyanagi *et al.*, 2002; Zhong and Lambowitz, 2003).

2.3. Reverse Splicing and Intron Mobility

As mentioned above, group II intron RNA acts as a self-splicing ribozyme to excise itself from mRNA. The splicing reaction can also occur in the reverse direction with the excised intron RNA re-inserting back into the mRNA molecule *in vitro* (Morl and Schmelzer, 1990). This ability to reverse splice turns out to be a key event in the mobility of group II introns into new locations, except in this case the intron RNA reverse splices directly into a DNA molecule rather than an RNA molecule.

Mobility of group II introns in bacteria is divided into two types of insertion events. The first is a highly efficient, site-specific insertion of the intron into what is called the homing site for the intron. The homing site is simply an analogous DNA sequence lacking the intron element. This type of mobility event is called retrohoming and does not involve homologous recombination (Cousineau *et al.*, 1998). In rare cases, apparently, a group II intron may insert into an unrelated or non-allelic DNA site in a process called retrotransposition (Cousineau *et al.*, 2000). In both cases reverse splicing and reverse transcription are key events in the mobilization of the intron to a new site.

From the limited information derived from sequenced bacterial genomes, group II introns appear to be the most common of the RT encoding elements found in the prokaryotes. About one fourth of all the bacterial genomes sequenced appear to contain at least one group II intron. Group II introns appear rarely among the few archaeal genomes that have been sequenced. An analysis of the insertion sites of group II introns in bacteria indicates that these intron DNAs only occasionally interrupt an ORF that requires the subsequent splicing of the intron to produce a functional gene product (Dai and Zimmerly, 2003). More commonly, group II introns are found associated with some kind of mobile DNA such as plasmids, IS elements, or transposons. When they are found in the chromosome, they commonly insert into intergenic DNA between two intact genes (Dai and Zimmerly, 2003).

Although the group II introns appear to be fairly common in prokaryotic genomes, they were not the first type of RT containing element to be discovered in bacteria. That distinction goes to the next type of retroelement found in the prokaryotes, the retron elements.

3. RETRONS AND msDNA

These strange retroelements, whose function is essentially unknown, have been observed to do only one thing; they synthesize large quantities of multicopy single-stranded DNA (msDNA). Indeed, it was the discovery of this unusual satellite msDNA that led to the first discovery of RT from a prokaryotic organism (Lampson *et al*., 1989b; Lim and Maas, 1989).

3.1. Organization of Retron DNA

Retrons are distinct DNA sequences of about 2000 base pairs that are found inserted into the bacterial chromosome or as part of a prophage DNA element. The functional organization of the retron is based on the only known activity of these retroelements, that is, the production of msDNA. Other possible functions like transposition are not known.

All retron elements have at least one large ORF designated *ret* (Fig. 3). The amino acid sequence of this ORF product, which can range from 298 to 586 amino acids, reveals an easily recognizable RT that is similar to eukaryotic RTs (Lampson *et al*., 1989b, 2002). Experiments with mutational variants, containing a deleted or inactivated ORF showed that these retrons fail to produce msDNA (Lampson *et al*., 1989b). Thus, the synthesis of msDNA requires a functional RT and is the first example of a reverse transcribed cDNA in bacteria. Immediately upstream to the start of the ORF are two genes designated *msr* and *msd*. These two genes are also required to synthesize msDNA and are situated in opposite directions such that their 3'ends overlap by several bases (Fig. 3 and Fig. 5). The genes *msr, msd* and *ret* are transcribed together as one mRNA and appear to be under the control of a promoter found within the retron DNA just upstream of *msr* (Herzer *et al*., 1992). Because msDNA is not an autonomously replicating satellite DNA, the three genes *msr, msd*, and *ret* essentially form an operon to synthesize msDNA. Some retrons may code for a second ORF but its function remains unknown (Sun *et al*., 1991).

3.2. The Retron RT

As noted for the intron protein of bacterial group II introns, the retron RT also contains the seven conserved domains 1-7 that roughly correspond to the fingers and palm region in the crystal structure of HIV RT (Fig. 3). Also like the intron protein, the retron RT contains the spacer regions 2a, 3a, and a very small 4a. The function of these spacer regions is not well understood. Beyond the RT domain at the C-terminus of the retron RT is a unique region of amino acids not found in other RTs, and sometimes designated domain "Y" (Fig. 3). The circular dichroism spectrum of this region in retron RT-Ec86 indicates mostly alpha-helical structures. Thus, this C-terminal domain may correspond to the three alpha-helicies (α-H, α-I, and α-J) of the thumb region in the crystal structure of HIV RT

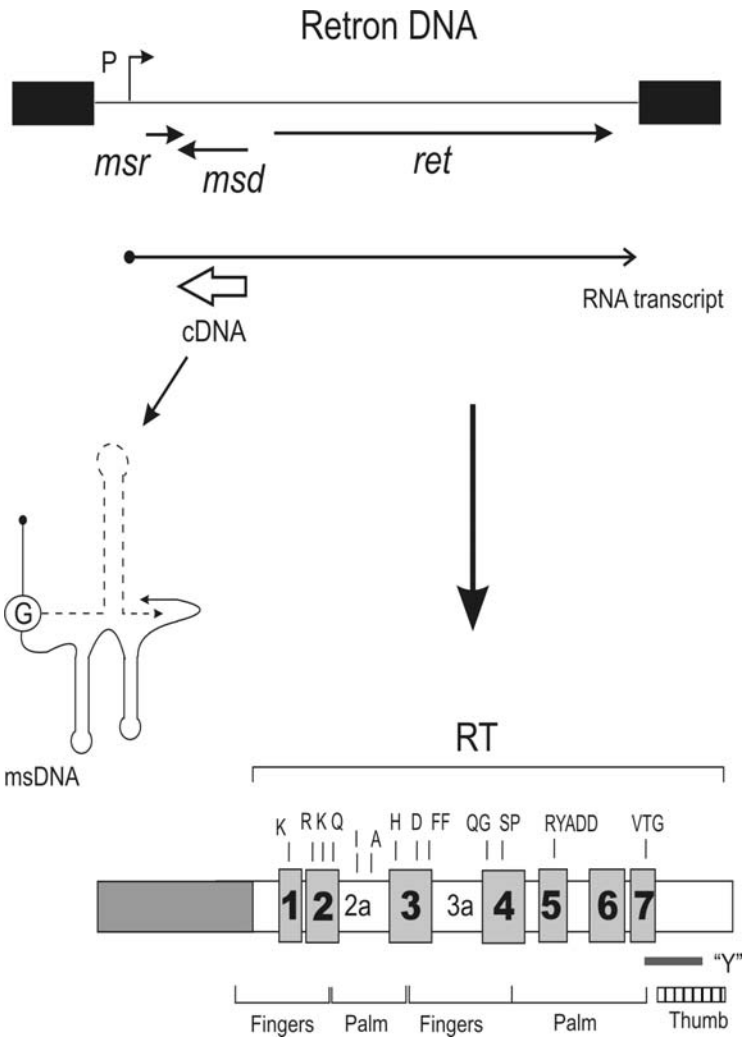


Figure 3. Retron DNA: an operon for making msDNA. Retron elements, composed of about 2 Kb of unique DNA (thin line), are usually inserted into the chromosome (black boxes) or into prophage DNA. Most retron elements contain a single ORF encoding RT (*ret*). Adjacent to *ret* are two genes designated *msr* and *msd*. The *msr* sequence corresponds to the RNA molecule in msDNA and the *msd* sequence corresponds to the DNA chain in msDNA. A cDNA copy of the *msd* gene produced by the retron RT leads to production of the satellite DNA known as msDNA. Like all RTs, the retron protein (bottom diagram) contains the seven blocks of conserved amino acids (boxes labelled 1-7) that fold into the fingers-palm region that is observed in the crystal structure of HIV RT (labelled brackets at the bottom). Some of the more highly conserved amino acids in this region of the retron RT are shown as single-letter abbreviations above the boxes. Based on circular dichroism analysis, the C-terminal part of the retron protein appears to have a structure similar to the thumb region in the crystal structure of HIV RT. This region of the protein binds to a specific structure in the template-primer RNA used to make msDNA. Some retron RTs have an extended N-terminal region of unknown function (shaded rectangle)

(Yamanaka *et al.*, 2002; Inouye *et al.*, 2004). Experiments with just the purified 66 amino acid C-terminal domain of retron RT-Ec86 indicates that this thumb region recognizes and binds to a specific stem-loop structure in the RNA template molecule used to synthesize msDNA. As discussed below, this is a critical step in the unusual priming mechanism used to initiate the synthesis of msDNA (Inouye *et al.*, 1999, 2004). It is interesting to note the similarity between the C-terminal thumb region of retron RT and the maturase (domain "X") of group II intron proteins. In both proteins this thumb structure appears to recognize and bind to a specific stem-loop structure in the RNA transcript encoding the RT. In the case of the intron protein this binding helps to maintain the folded structure of the intron RNA for splicing. In the case of the retron protein this binding helps to maintain the structure needed to prime the start of synthesis of msDNA (see below).

Some retron RTs have an extremely long N-terminal extension beyond the RT domains 1 through 7. For example, the RT from retron Mx162, from the myxobacterium *Myxococcus xanthus*, contains a 165 amino acid extension before the RT domain 1 (Inouye *et al*, 1989). The function of this extended region is unknown.

3.3. The Unique Structure of msDNA

msDNA is a strange satellite DNA first discovered in *M. xanthus* (Yee *et al.*, 1984). In the case of msDNA-Ne144 from the related myxobacterium *Nannocystis exedens*, the satellite DNA is composed of a 144 nucleotide single-stranded DNA that folds into a stable stem-loop structure. Covalently linked to the 5' end of the DNA chain is a 72 base, single-strand of RNA. However, the RNA strand retains both a free 5' end and a free 3' end and is instead joined to the DNA strand at an internal guanosine nucleotide via a 2'–5' phosphodiester bond (circled G, Fig. 4). msDNA is thus a unique molecule in that a DNA molecule is joined to an RNA molecule by a 2'–5' phosphodiester linkage (Dhundale *et al.*, 1987; Furuichi *et al.*, 1987). In addition to RNA and DNA, msDNA is most likely bound in a complex with protein. The major protein in this DNA-RNA-protein complex appears to be the retron RT (Lampson *et al.*, 1990), although other host proteins may also bind to msDNA. For example, RNase H produced by the host cell may bind to msDNA, at least initially, since it appears to be involved in the synthesis of msDNA (Lima and Lim, 1995; Shimamoto *et al*, 1995).

Different retron elements produce distinct msDNAs with generally little similarity in the primary nucleotide sequence of either the DNA molecule or the RNA molecule. All msDNAs, however, appear to contain, at least initially, the basic secondary structure shown in Fig. 4. Some msDNAs, apparently, undergo some modification after they are synthesized. For example, the entire RNA strand including the 2'–5' branch linkage is removed from msDNA-Ec83 from *Escherichia coli* resulting in a mature msDNA that is only composed of a single-stranded DNA molecule (Lim, 1992; Kim *et al.*, 1997). The unusual structure of msDNA appears to be a consequence of the way this DNA is synthesized (see below). Why some msDNAs are then further modified is unknown.

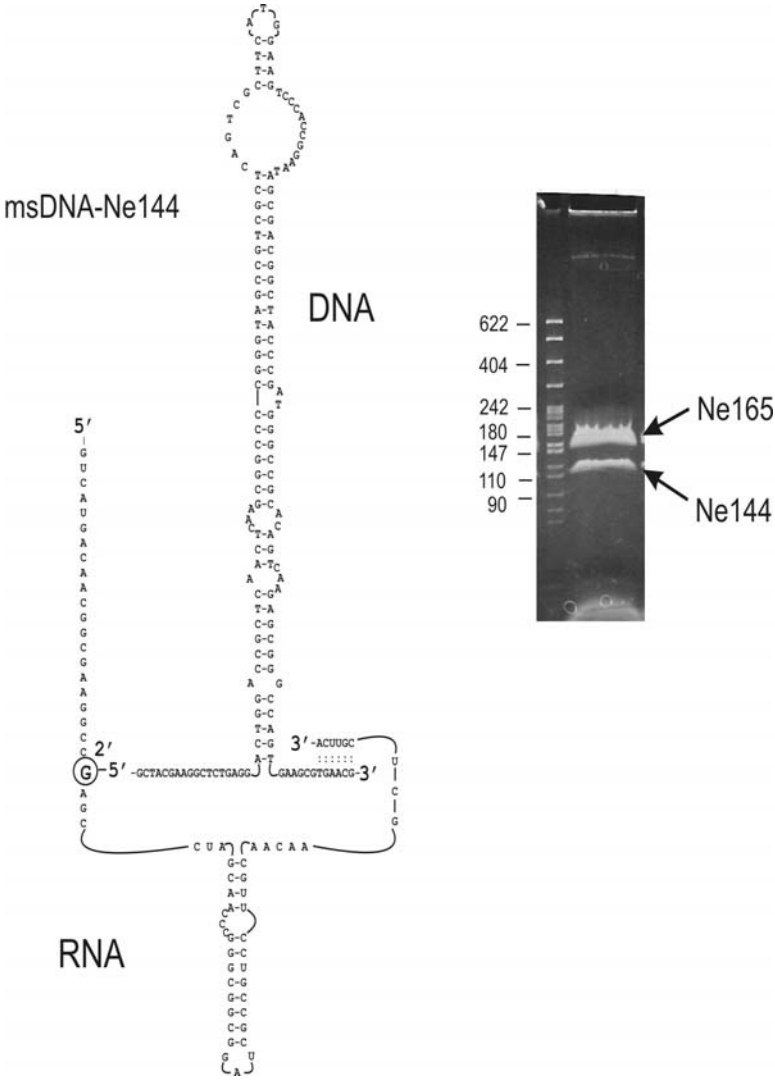


Figure 4. The unique structure of msDNA. msDNA is a small, abundant satellite DNA easily observed by gel electrophoresis of DNA extracts. For example, *N. exedens* produces two different msDNAs (inset photo) designated Ne165 and Ne144 based on the size of the DNA strand in each molecule. The structure of msDNA-Ne144 is shown. This msDNA is composed of a single-strand of DNA containing 144 deoxyribonucleotides that folds into a long stable stem-loop structure. Joined to the 5'-end of the DNA chain is a 72 ribonucleotide, single-stranded RNA molecule. The RNA strand is joined to the 5'-end of the DNA molecule via a 2'–5' phosphodiester bond at a specific internal guanosine residue (circled G) of the RNA molecule. In addition to stem-loop folding in the RNA chain, a small region of base-pairing occurs between the very 3'-end of the DNA chain and the 3'-end of the RNA chain

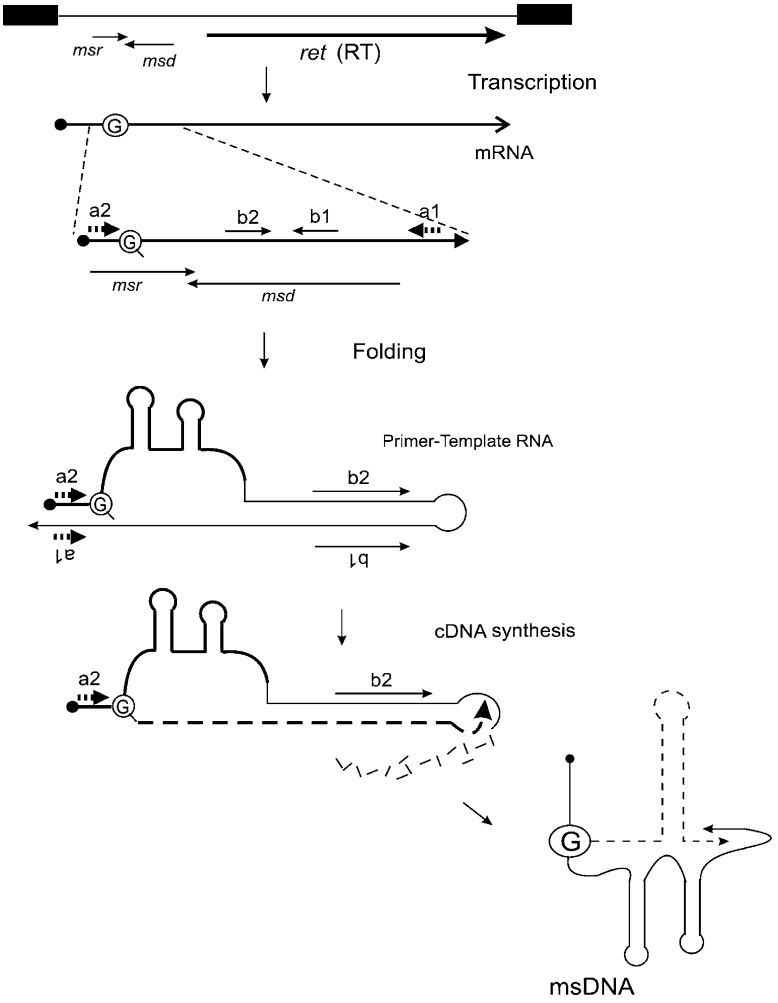


Figure 5. msDNA is synthesized by reverse transcription. A smaller region of a longer mRNA transcript of the retron *msr-msd-ret* operon will serve as both a primer and template for the synthesis of msDNA. Folding of the template RNA (mediated by inverted repeats a1-a2 and b1-b2) is crucial to form the 2'-OH primer at a specific guanosine residue (circled G) to initiate cDNA synthesis by the retron RT. After synthesis of cDNA (dashed line) is complete, part of the RNA template remains joined to the 5'-end of the cDNA molecule to yield the complete msDNA

3.4. msDNA is Synthesized by Reverse Transcription

The interesting way msDNA is reverse transcribed helps explain its unusual structure. Synthesis of msDNA begins from an RNA transcript (of the *msr-msd* genes) that folds into a specific secondary structure (Dhundale *et al.*, 1987). The 2'-OH group of a specific guanosine (circled G, Fig. 5) within this folded RNA

serves as a primer to initiate cDNA synthesis (Hsu *et al.*, 1989). The retron RT utilizes this 2'-OH group to incorporate the first deoxyribonucleotide using the folded RNA as a template. In close concert with the extension of the cDNA strand is removal of the lagging RNA template by an RNase H activity apparently provided by the host cell (Lampson *et al*, 1989a; Shimamoto *et al*, 1995). Host RNase H plays a similar role in removal of the RNA template in group II intron retro-mobility (Smith *et al.*, 2005). At a specific location on the RNA template, cDNA polymerization stops, resulting in the msDNA molecule observed in cell extracts. That is, a short cDNA chain covalently joined to the remainder of the RNA molecule that served as a primer-template for its synthesis.

3.5. Retrons and the Host Cell

While group II introns occur quite frequently in bacterial genomes, retron elements appear far less frequently and tend to occur sporadically within a population of the same species. For example, among the 75 strains of *E. coli* that make up the ECOR collection, all 75 strains contain at least one group II intron element and one strain, ECOR 38, has as many as 15 copies of the same intron element in its genome (Dai and Zimmerly, 2003). In contrast, only 11 of the 75 ECOR strains contains a retron element and no report exists of a bacterial chromosome containing more than one copy of the same retron element (Herzer *et al.*, 1990). Nevertheless, retrons are widely found among diverse taxonomic groups of the prokaryotes including the alpha, beta, gamma, and delta proteobacteria, fusobacteria, Gram positives, cyanobacteria, and one example from the archeon *Methanosarcina acetivorans* (Rest and Mindell, 2003; Lampson *et al*, 2005).

Although it is not known if retrons are mobile DNAs, an increase in the frequency of spontaneous mutations is observed when thousands of copies of the msDNA molecule are produced during over-expression of a retron in *E. coli* (Maas *et al*, 1994, 1996). Here, an increased frequency of frame-shift mutations occurred due to binding and titrating out of the host repair protein MutS to mis-matched base pairs in msDNA. In another interesting observation, numerous partial copies of the msDNA sequence (*i.e.* the *msd* gene) were detected in a bacterial chromosome (Lampson and Rice, 1997; Lampson *et al.*, 2002). Evidence suggested that these repeated DNA sequences could be retrotransposed copies of msDNA produced by the corresponding retron RT.

4. DIVERSITY GENERATING ELEMENTS

The third type of prokaryotic RT containing element is in the recently described diversity generating retroelements. These elements are so named because they appear to function as a special mechanism to produce sequence variation in a phage gene encoding a tail protein. By generating nucleotide substitutions at 23 sites in a 134 base-pair VR region at the 3' end of the phage gene, many variants of the phage tail protein can be produced (Fig. 6) (Liu *et al.*, 2002). By changing certain amino acids at the C-terminal end of this tail protein, the phage can recognize and

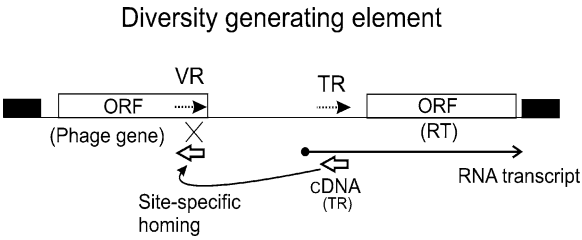


Figure 6. Diversity generating elements (DGEs) produce DNA sequence variation by reverse transcription. DGEs contain an ORF encoding a typical bacterial RT. Upstream of the RT-ORF is a matching pair of directly repeated sequences designated VR and TR. Synthesis of a cDNA copy of the TR sequence by the DGE RT eventually introduces nucleotide variation in the VR sequence. Because the VR repeat sequence is part of an upstream ORF, this produces changes in the amino acid sequence of the corresponding protein product (in this example variation occurs in a phage tail protein). See text for more detail

bind to alternative receptor molecules on the surface of the bacterial host that it infects. Thus, the presence of this diversity generating retroelement can confer a direct selective advantage on the phage organism (Doulatov *et al.*, 2004).

The diversity generating elements appear to be composed of a long ORF encoding a typical RT similar to retroviral RT and especially analogous to the bacterial RTs. The protein product of this ORF has been shown to be enzymatically active in a standard RT assay (Liu *et al.*, 2002). Upstream from the RT-ORF DNA is a pair of 134-base, directly repeated sequences designated VR and TR. The TR sequence is just 5' to the start of the RT-ORF and is termed the template sequence. Further upstream, the VR sequence is found as part of the extreme 3' end of an upstream ORF (Fig. 6). The VR sequence is termed the variable region because nucleotides at 23 specific positions in this sequence will change among variants produced by this element (Liu *et al.*, 2002, 2004). With the TR region of a mRNA serving as a template, sequence variation is thought to be produced by the reverse transcription of a cDNA copy of the TR sequence (Fig. 6). This cDNA copy then inserts into the VR region by what is termed a site-specific homing event (Doulatov *et al.*, 2004). Eventually the cDNA copy containing the nucleotide substitutions replaces the original VR sequence. This, then produces amino acid changes in the phage tail protein encoded by this upstream ORF.

Through sequence homology searches of sequenced microbial genomes, diversity generating elements have so far been found in several diverse groups of bacteria including *Bordetella, Vibrio, Bacteroides, Treponema*, and several cyanobacteria (Doulatov *et al.*, 2004).

5. APPLIED USES FOR PROKARYOTIC RTS

Currently, almost all commercially marketed RTs are eukaryotic enzymes. That is, they are purified or cloned from animal retroviruses or are genetically engineered derivatives of these enzymes. For example, the mutational inactivation of the

C-terminal RNase H domain from the RT of Moloney murine leukemia virus (M-MLV) yields an RT with greatly improved efficiency of cDNA synthesis *in vitro* (Kotewicz *et al.*, 1988). Retroviral RTs are today extensively used for many applications that require synthesis of cDNA, from construction of cDNA libraries to RT-PCR. However, the newly described prokaryotic RTs are a highly diverse group of enzymes with novel properties and represent an emerging new tool with potential new applications for RTs.

5.1. Targetrons

An example of a bacterial RT that is currently marketed as a genetic tool is the so called targetron (Sigma-Aldrich). Targetrons take advantage of the way the group II intron RT works in conjunction with intron RNA to mediate insertion of the intron into a site on DNA (Guo *et al.*, 2000). While the intron RT protein binds and unwinds the target DNA molecule, it is specific sequences within the intron RNA that recognize and base-pair with the homing site (insertion site) in DNA. This allows for a very efficient and site-specific insertion of the intron into the homing site on the target DNA molecule by the retrohoming mechanism. The targetron system allows the user to reprogram the group II intron RNA molecule to recognize a new insertion site in a DNA sequence of their choosing. This is done by changing specific nucleotides in the intron RNA, using a mutagenic primer, so that it will recognize and base pair to a new homing site (Guo *et al.*, 2000). The resulting targetron system is a highly efficient, site-specific insertion vector that can insert and disrupt almost any desired bacterial gene in a cell where this group II intron can be expressed (Frazier *et al.*, 2003; Zhong *et al.*, 2003). In addition, by using a mutagenic primer-PCR reaction that incorporates random nucleotides, the intron RNA can be reprogrammed to insert into many random sites in a target chromosome. Such a randomized targetron vector was used to produce a gene "knock-out" library in *E. coli* where most non-essential genes were shown to be disrupted by an intron insertion (Yao *et al.*, 2005).

5.2. Thermophilic RT

Recently, a group II intron was discovered in the thermophilic bacterium *Geobacillus stearothermophilus*. The intron encoded protein was cloned, over-expressed in *E. coli* and shown, like its host organism, to be heat stable. The RT can reverse transcribe RNA at temperatures as high as 75 °C (Vellore *et al.*, 2004). One of the big problems in using retroviral RTs to produce cDNA is that the enzyme will prematurely pause during polymerization of DNA. This is due to formation of secondary structures in the RNA template resulting in truncated cDNA products (Harrison *et al.*, 1998). The length of cDNAs produced by RT can be greatly increased by carrying out reverse transcription at temperatures above 55 °C. These hot temperatures will melt the secondary folding in the RNA template and alleviate pausing of the RT (Fu and Stuve, 2003; Hawkins *et al.*, 2003). Several

retroviral RTs have now been modified to increase their thermo-stability and activity at higher temperatures for this reason. The discovery of a genuine heat-stable RT from a thermophilic bacterium that remains active at higher temperatures could also prove useful for synthesizing long cDNAs. A group II intron has also been reported from the related thermophile *G. kaustophilus* (Chee and Takami, 2005). This intron interrupts a functional *recA* gene and was shown to splice at temperatures above 70 °C *in vivo*.

5.3. msDNA

Although its function remains elusive, a number of interesting potential applications for the retron-msDNA system have been explored. For example, msDNA can be used as a type of antisense DNA to repress the expression of a particular gene. In the stem-loop structure of msDNA the unpaired bases in the loop can be greatly expanded to include more DNA without affecting the high level of msDNA produced by the retron RT. Indeed, such an engineered retron was made so that its msDNA contained a large loop of DNA sequence complementary to the translation initiation region on the transcript for the *lpp* gene of *E. coli* (Mao *et al.*, 1995). The *lpp* gene codes for the highly abundant *E. coli* lipoprotein of the outer membrane. Over-expression of this engineered retron caused a 77% reduction in the level of lipoprotein in these cells. This was due to the resulting production of thousands of copies of this modified msDNA molecule which apparently acts as an antisense DNA *in vivo* to reduce the expression of the *lpp* gene. Other potential uses of the versatile retron-msDNA system can be envisioned that have not been explored experimentally. This might include a system for producing cDNA *in vivo*, or over-production of a competing DNA binding site for a regulatory protein (Lampson *et al.*, 2001).

ACKNOWLEDGMENTS

I wish to thank Masayori Inouye, Sumiko Inouye, and Marlene Belfort for very helpful comments and suggestions.

REFERENCES

Belfort, M. and Roberts, R.J. (1997). Homing endonucleases: keeping the house in order. Nucleic Acid Res. *25*, 3379–3388.

Blocker, F.J.H., Mohr, G., Conlan, L.H., Qi, L., Belfort, M. and Lambowitz, A.M. (2005). Domain structure and three-dimensional model of a group II intron-encoded reverse transcriptase. RNA *11*, 14–28.

Chee, G.-J. and Takami, H. (2005). Housekeeping *recA* gene interrupted by group II intron in the thermophilic *Geobacillus kaustophilus.* Gene (in press).

Cousineau, B., Smith, D., Lawrence-Cavanagh, S., Mueller, J.E., Yang J., Mills D., Nanias, D.A., Dunny, G.M., Lambowitz, A.M. and Belfort, M. (1998). Retrohoming of a bacterial group II intron, mobility via complete reverse splicing, independent of homologous DNA recombination. Cell *94*, 451–462.

Cousineau, B., Lawrence, S., Smith, D. and Belfort, M. (2000). Retrotransposition of a bacterial group II intron. Nature *404*, 1018–1021.

Dai, L. and Zimmerly, S. (2003). Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior. Nucleic Acid Res. *30*, 1091–1102.

Dhundale, A., Lampson, B., Furuichi, T., Inouye, M. and Inouye, S. (1987). Structure of msDNA from *Myxococcus xanthus*: evidence for a long self-annealing RNA precursor for the covalently linked, branched RNA. Cell *51*, 1105–1112.

Doulatov, S., Hodes, A., Dai, L., Mandhana, N., Liu, M., Deora, R., Simons, R.W., Zimmerly, S. and Miller, J.F. (2004). Tropism switching in *Bordetella* bacteriophage defines a family of diversity-generating retroelements. Nature *431*, 476–480.

Eickbush, T.H. (1994). Origin and evolutionary relationships of retroelements. In The evolutionalry biology of viruses, S.S. Morse, ed. New York, NY: Raven Press. pp. 127–157.

Eickbush, T.H. and Malik, H.S. (2002). Origins and evolution of retrotransposons. In Mobile DNA II, N.L. Craig *et al*., eds. Washington, DC: ASM Press. pp. 1111–1144.

Ferat, J.L. and Michel, F. (1993). Group II self-splicing introns in bacteria. Nature *364*, 358–361.

Frazier, C.L., San Filippo, J., Lambowitz, A. M and Mills, D.A. (2003). Genetic manipulation of *Lactococcus lactis* by using targeted group II introns: generation of stable insertions without selection. Appl. Environ. Microbiol. *69*, 1121–1128.

Fu, G.K. and Stuve, L.L. (2003). Improved method for the construction of full-length enriched cDNA libraries. BioTechniques *34*, 954–957.

Furuichi, T., Dhundale, A., Inouye, M. and Inouye, S. (1987). Branched RNA covalently linked to the 5' end of a single-stranded DNA in *Stigmatella aurantiaca*: structure of msDNA. Cell *48*, 47–53.

Guo, H., Karberg, M., Long, M., Jones, J.P., Sullenger, B. and Lambowitz, A.M. (2000). Group II introns designed to insert into therapeutically relevant DNA target sites in human cells. Science *289*, 452–457.

Guo, H., Zimmerly, S., Perlman, P.S. and Lambowitz, A.M. (1997). Group II intron endonucleases use both RNA and protein subunits for recognition of specific sequences in double-stranded DNA. EMBO J. *16*, 6835–6848.

Harrison, G.P., Mayo, M.S., Hunter, E. and Lever, A.M.L. (1998). Pausing of reverse transcriptase on retroviral RNA templates is influenced by secondary structures both 5' and 3' of the catalytic site. Nucleic Acid Res. *26*, 3433–3442.

Hawkins, P.R., Jin, P. and Fu, G.K. (2003). Full-length cDNA synthesis for long-distance RT-PCR of large mRNA transcripts. BioTechniques *34*, 768–773.

Herzer, P.J., Inouye, S., Inouye, M. and Whittam, T.S. (1990). Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. J. Bacteriol. *172*, 6175–6181.

Herzer, P.J., Inouye, S. and Inouye, M. (1992). Retron-Ec107 is inserted into the *Escherichia coli* genome by replacing a palindromic 34 bp intergenic sequence. Mol. Microbiol. *6*, 345–354.

Hsu, M-Y., Inouye, S. and Inouye, M. (1989). Structural requirements of the RNA precursor for the biosynthesis of the branched RNA-linked msDNA of *Myxococcus xanthus*. J. Biol. Chem. *264*, 6214–6219.

Ichiyanagi, K., Beauregard, A., Lawrence, S., Smith, D., Cousineau, B. and Belfort, M. (2002). Retrotransposition of the Ll.LtrB group II intron proceeds predominantly via reverse splicing into DNA targets. Mol. Microbiol. *46*, 1259–1272.

Inouye, M., Ke, H., Yashio, A., Yamanaka, K., Nariya, H., Shimamoto, T. and Inouye, S. (2004). Complex formation between a putative 66-residue thumb domain of bacterial reverse transcriptase RT-Ec86 and the primer recognition RNA. J. Biol. Chem. *279*, 50735–50742.

Inouye, S., Hsu, M-Y., Eagle, S. and Inouye, M. (1989). Reverse transcriptase associated with the biosynthesis of the branched RNA-linked msDNA in *Myxococcus xanthus*. Cell *56*, 709–717.

Inouye, S., Hsu, M-Y., Xu, A. and Inouye, M. (1999). Highly specific recognition of primer RNA structures for 2′-OH priming reaction by bacterial reverse transcriptases. J. Biol. Chem. *274*, 31236–31244.

Kazazian, H.H. (2004). Mobile elements: drivers of genome evolution. Science *303*, 1626–1632.

Kim, K., Jeong, D. and Lim, D. (1997). A mutational study of the site-specific cleavage of Ec83, a multicopy single-stranded DNA (msDNA): nucleotides at the msDNA stem are important for its cleavage. J. Bacteriol. *179*, 6518–6521.

Kotewicz, M.L., Sampson, C.M., D'Alessio, J.M. and Gerard, G. F. (1988). Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H. Nucleic Acid Res. *16*, 265–277.

Lambowitz, A.M. and Zimmerly, S. (2004). Mobile group II introns. Annu. Rev. Genet. *38*, 1–35.

Lampson, B.C., Inouye, M. and Inouye, S. (1989a). Reverse transcriptase with concomitant ribonuclease H activity in the cell-free synthesis of branched RNA-linked msDNA of *Myxococcus xanthus*. Cell *56*, 701–707.

Lampson, B.C., Inouye, M. and Inouye, S. (2001). The msDNAs of bacteria. Prog. Nucleic Acid Res. Mol. Biol. *67*, 65–91.

Lampson, B.C., Inouye, M. and Inouye, S. (2005). Retrons, msDNA, and the bacterial genome. Cytogenet. Genome Res. *110*, 491–499.

Lampson, B.C. and Rice, S.A. (1997). Repetitive sequences found in the chromosome of the myxobacterium *Nannocystis exedens* are similar to msDNA: a possible retro-transposition event in bacteria. Mol. Microbiol. *23*, 813–823.

Lampson, B.C., Sun, J., Hsu, M-Y., Vallejo-Ramirez, J., Inouye, S. and Inouye, M. (1989b). Reverse transcriptase in a clinical strain of *E. coli*: its requirements for production of branched RNA-linked msDNA. Science *243*, 1033–1038.

Lampson, B.C., Xu, C., Rice, S.A. and Inouye, S. (2002). A partial copy of msDNA from a new retron element is likely a retro-transposed DNA found in the myxobacterium *Nannocystis exedens*. Gene *299*, 251–261.

Lampson, B.C., Viswanathan, M., Inouye, M. and Inouye, S. (1990). Reverse transcriptase from *Escherichia coli* exists as a complex with msDNA and is able to synthesize double-stranded DNA. J. Biol. Chem. *265*, 8490–8496.

Lim, D. (1992). Structure and biosynthesis of unbranched multicopy single-stranded DNA by reverse transcriptase in a clinical *Escherichia coli* isolate. Mol. Microbiol. *6*, 3531–3542.

Lim, D. and Maas, W.K. (1989). Reverse transcriptase-dependent synthesis of a covalently linked, branched DNA-RNA compound in *E. coli* B. Cell *56*, 891–904.

Lima, T.M.O. and Lim, D. (1995). Isolation and characterization of host mutants defective in msDNA synthesis: role of ribonuclease H in msDNA synthesis. Plasmid *33*, 235–238.

Liu, M., Deora, R., Doulatov, S.R., Gingery, M., Eiserling, F. A., Preston, A., Maskell, D.J., Simons, R.W., Cotter, P.A., Parkhill, J. and Miller, J.F. (2002). Reverse transcriptase-mediated tropism switching in *Bordetella* bacteriophage. Science *295*, 2091–2094.

Liu, M., Gingery, M., Doulatov, S.R., Liu, Y., Hodes, A., Baker, S., Davis, P., Simmonds, M., Churcher, C., Mungall, K., Quail, M. A., Preston, A., Harvill, E.T., Maskell, D.J., Eiserling, F.A., Parkhill, J. and Miller, J.F. (2004). Genomic and genetic analysis of *Bordetella* bacteriophages encoding reverse transcriptase-mediated tropism-switching cassettes. J. Bacteriol. *186*, 1503–1517.

Lue, N.F. (2004). Adding to the ends: what makes telomerase processive and how important is it? Bioessays *26*, 955–962.

Maas, W.K., Wang, C., Lima, T., Zubay, G. and Lim, D. (1994). Multicopy single-stranded DNAs with mismatched base pairs are mutagenic in *Escherichia coli*. Mol. Microbiol. *14*, 437–441.

Maas, W.K., Wang, C., Lima, T., Hach, A. and Lim, D. (1996). Multicopy single-stranded DNA of *Escherichia coli* enhances mutation and recombination frequencies by titrating MutS protein. Mol. Microbiol. *19*, 505–509.

Mao, J.-R., Shimada, M., Inouye, S. and Inouye, M. (1995). Gene regulation by antisense DNA produced *in vivo*. J. Biol. Chem. *270*, 19684–19687.

Martinez-Abarca, F., Garcia-Rodriguez, F.M. and Toro, N. (2000). Homing of a bacterial group II intron with an intron-encoded protein lacking a recognizable endonuclease domain. Mol. Microbiol. *35*, 1405–1412.

Martinez-Abarca, F. and Toro, N. (2000). Group II introns in the bacterial world. Mol. Microbiol. *38*, 917–926.

Matsurra, M., Noah, J.W. and Lambowitz, A.M. (2001). Mechanism of maturase-promoted group II intron splicing. EMBO J. *20*, 7259–7270.

Michel, F. and Lang, B.F. (1985). Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses. Nature *316*, 641–643.

Mohr, G., Perlman, P.S. and Lambowitz, A.M. (1993). Evolutionary relationships among group II intron-encoded proteins and identification of a conserved domain that may be related to maturase function. Nucleic Acid Res. *21*, 4991–4997.

Morl, M. and Schmelzer, C. (1990). Integration of group II intron *bl*1 into a foreign RNA by reversal of the self-splicing reaction *in vitro*. Cell *60*, 629–636.

Rest, J.S. and Mindell, D.P. (2003). Retroids in Archaea: phylogeny and laterial origins. Mol. Biol. Evol. *20*, 1134–1142.

Saldanha, R., Mohr, G., Belfort, M. and Lambowitz, A.M. (1993). Group I and group II introns. FASEB J. *7*, 15–24.

Shimamoto, T., Shimada, M., Inouye, M. and Inouye, S. (1995). The role of ribonuclease H in multicopy single-stranded DNA synthesis in retron-Ec73 and retron-Ec107 of *Escherichia coli*. J. Bacteriol. *177*, 264–267.

Singh, R.N., Saldanha, R.J., D'Souza, L.M. and Lambowitz, A.M. (2002). Binding of a group II intron-encoded reverse transcriptase-maturase to its high affinity intron RNA binding site involves sequence-specific recognition and autoregulates translation. J. Mol. Biol. *318*, 287–303.

Smith, D., Zhong, J., Matsuura, M., Lambowitz, A.M. and Belfort, M. (2005). Recruitment of host functions suggests a repair pathway for late steps in group II intron retrohoming. Genes Dev. *19*, 2477–2487.

Steitz, T.A. (1999). DNA polymerases: structural diversity and common mechanisms. J. Biol. Chem. *274*, 17395–17398.

Sun, J., Inouye, M. and Inouye, S. (1991). Association of a retroelement with a P4-like cryptic prophage (retron-phage φ R73) integrated into the selenocystyl tRNA gene of *Escherichia coli*. J. Bacteriol. *173*, 4171–4181.

Toor, N., Hausner, G. and Zimmerly, S. (2001). Co-evolution of group II intron RNA structures with their intron-encoded reverse transcriptases. RNA *7*, 1142–1152.

Vellore, J., Moretz, S.E. and Lampson, B.C. (2004). A group II intron-type open reading frame from the thermophile *Bacillus (Geobacillus) stearothermophilus* encodes a heat-stable reverse transcriptase. Appl. Environ. Microbiol. 70, 7140–7147.

Xiong, Y. and Eickbush, T.H. (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. *9*, 3353–3362.

Yamanaka, K., Shimamoto, T., Inouye, S. and Inouye, M. (2002). Retrons. In Mobile DNA II, N.L. Craig *et al*., eds. Washington, DC: ASM Press. pp. 784–795.

Yao, J., Zhong, J. and Lambowitz, A.M. (2005). Gene targeting using randomly inserted group II introns (targetrons) recovered from an *Escherichia coli* gene disruption library. Nucleic Acid Res. *33*, 3351–3362.

Yee, T., Furuichi, T., Inouye, S. and Inouye, M. (1984). Multicopy single-stranded DNA isolated from a Gram-negative bacterium, *Myxococcus xanthus*. Cell *38*, 203–209.

Zhong, J., Karberg, M. and Lambowitz, A.M. (2003). Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. Nucleic Acid Res. *31*, 1656–1664.

Zhong, J. and Lambowitz, A.M. (2003). Group II intron mobility using nascent strands at DNA replication forks to prime reverse transcription. EMBO J. *22*, 4555–4565.

Zimmerly, S., Hausner, G. and Wu, X.-C. (2001). Phylogenetic relationships among group II intron ORFs. Nucleic Acid Res. *29*, 1238–1250.

CHAPTER 24

DICER: STRUCTURE, FUNCTION AND ROLE IN RNA-DEPENDENT GENE-SILENCING PATHWAYS

JUSTIN M. PARE[1] AND TOM C. HOBMAN[1,2*]

[1]*Departments of Cell Biology and*
[2]*Medical Microbiology and Immunology, University of Alberta, Edmonton, AB, Canada*
[*]*tom.hobman@ualberta.ca*

1. INTRODUCTION

Dicer enzymes are double strand (ds)-specific RNases whose cleavage products are used to provide specificity in RNA-based gene-silencing pathways. They are encoded by the genomes of all multicellular and most unicellular eukaryotes. In mammals, Dicer is essential for viability. Evidence suggests that these modular 200 kD enzymes function in the cytoplasm in conjunction with Dicer-binding proteins to generate and load small RNAs onto ribonucleoprotein complexes that effect post-transcriptional gene-silencing. Dicer activity is also required for some types of transcriptional gene-silencing in the nucleus. The importance of Dicer as a central regulator of gene expression is illustrated by recent estimates that the expression of as many as one third of all human genes are regulated by Dicer cleavage products (Lewis *et al*, 2005). In this chapter, we discuss the structure and function of Dicer and related enzymes in RNA-based gene-silencing pathways.

2. STRUCTURE AND FUNCTION

2.1. RNase III Family Members

Dicer is a core member of the RNA interference (RNAi) apparatus that effects gene-silencing at transcriptional and post-transcriptional levels. It belongs to a large family of ribonucleases known as the RNase III superfamily. There are three subfamilies within this superfamily, each represented by their type members: (1) *Escherichia* coli RNase III, (2) Drosha and (3) Dicer. The subfamilies differ considerably with respect to their structural complexity and functional roles within the cell.

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 421–438

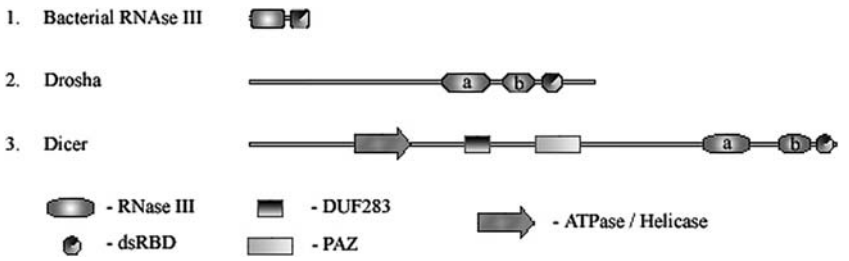


Figure 1. Domain organization of RNase III superfamily members

The members of the *E. coli* RNase III subfamily are relatively small (150–250 amino acid residues) and are structurally the simplest members of the superfamily. They are composed of a single RNase III domain together with a double stranded RNA binding domain (dsRBD) (Fig. 1). By comparison, Drosha subfamily members are structurally more complex than the bacterial RNase III isoforms. Members of this group contain two RNase III domains (a and b), a dsRBD and relatively long amino-termini (Fig. 1). The latter segment contains a proline-rich region and an arginine/serine-rich (RS-rich) region, both of which are thought to participate in protein-protein interactions: (Fortin *et al*., 2002; Wu *et al*., 2000).

The third class of RNase III enzymes is represented by Dicer family members. Similar to Drosha, Dicer enzymes contain two RNase III domains and a dsRBD, but lack proline-rich and RS-rich domains (Fig. 1). They are the only members of the RNase III superfamily that contain PAZ and RNA helicase domains. In addition, most Dicer isoforms possess a centrally located *d*omain of *u*nknown *f*unction (DUF283).

2.2. RNase III

True to its function as a ribonuclease, RNase III was originally isolated from *E. coli* as a dsRNA degrading activity (Robertson *et al*., 1968). Subsequently, it was determined that members of this subfamily are important for the maturation of ribosomal RNA, transfers RNAs and mRNAs (Bram *et al*., 1980; Nicholson, 1996; Regnier and Grunberg-Manago, 1990; Steege, 2000; Venema and Tollervey, 1999). In some organisms, RNase III enzymes are also involved in rRNA fragmentation (Bram *et al*., 1980; Gegenheimer and Apirion, 1980; Gegenheimer and Apirion, 1981). This process entails the removal of short nucleotide stretches from the mature domain of rRNA molecules without subsequent re-ligation of the cleaved fragments (Gerbi, 1995; Gray, 1995). For example, the 16S and 23S rRNA subunits are initially transcribed as a single primary transcript, which folds into a hairpin structure. The hairpin is cleaved by RNase III to generate the mature 16S and 23S rRNA subunits. In addition to their roles in rRNA processing, RNase III enzymes may also have regulatory functions that are not dependent upon catalytic activity. Specifically, the RNase III homolog, Rnt1, is required for normal cell cycle progression in budding yeast (Catala *et al*., 2004). In this case, the RNase activity of RntI is not required for

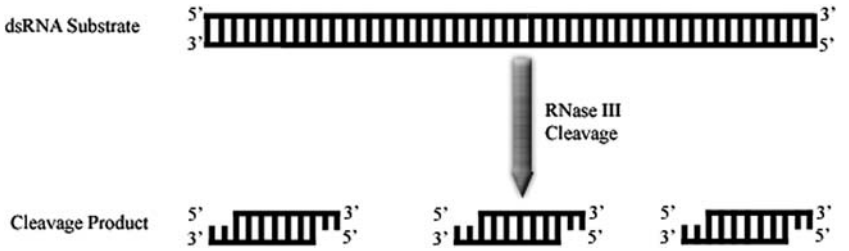


Figure 2. RNase III cleavage. Cleavage of long dsRNA substrates by RNase III enzymes yields uniformly sized dsRNA cleavage products with 2 nt 3' overhangs and 5' phosphate groups

its role in cell cycle regulation. Finally, it has been reported that RNase III enzymes can bind to certain mRNA substrates, a scenario which is at least consistent with the potential for these enzymes to affect translation directly or indirectly by decreasing the stability of the mRNAs (Dasgupta *et al*., 1998).

Activity of bacterial RNase III enzymes requires the formation of homodimers (Dunn, 1976; Robertson *et al*., 1968). Dimerization of two RNase III molecules results in the creation of a catalytic center comprised of two opposing RNase III domains. As mentioned above, this arrangement results in the positioning of the two catalytic sites in such a way that both strands of the dsRNA are cleaved. All members of the RNase III superfamily generate characteristic cleavage products that contain 5' phosphate groups and 2 nt 3' overhangs (Fig. 2) that terminate with hydroxyl groups (Nicholson, 1996). Although the lengths of cleavage products vary depending upon the particular enzyme, the products are of uniform sizes. For example, bacterial RNase III enzymes generate RNA duplexes that are approximately 11 nt long (Nicholson, 1996). In contrast, depending upon the isoform, Dicer cleavage products are duplexes which vary in length from 20 to 25 nt.

2.3. Drosha

Drosha family members localize predominantly to the nucleus and like their Class I relatives, are important for processing of pre-ribosomal precursors. Originally termed human RNase III, Drosha was recognized as an essential enzyme from experiments which showed that inhibiting the expression of this enzyme resulted in cell death (Wu *et al*., 2000). Subsequent work by the Kim laboratory revealed that Drosha activity is required upstream of Dicer in the microRNA (miRNA) biogenesis pathway (Lee *et al*., 2003). Specifically, this enzyme is responsible for cleaving the long primary-miRNAs (pri-miRNAs) into shorter pre-miRNAs (Lee *et al.*, 2003) (Fig. 3). Pri-miRNAs are either transcribed as independent transcripts or are derived from the intronic regions of protein-encoding genes. After they are processed by Drosha, pre-miRNAs are exported from the nucleus by an Exportin-5-dependent pathway (Bohnsack *et al*., 2004; Lund *et al*., 2004; Yi *et al*., 2003). This export factor recognizes the 2 nt 3' overhangs on Drosha cleavage products.

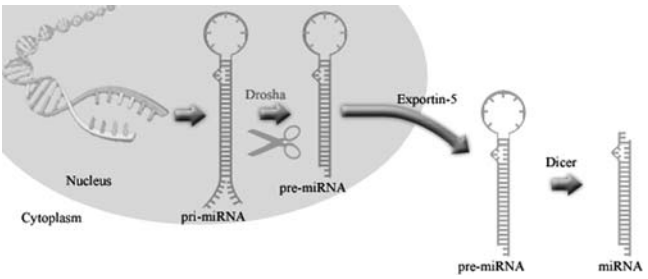


Figure 3. Maturation of miRNAs by Drosha and Dicer. Cleavage of pri-miRNA in the nucleus by Drosha leads to the formation of pre-miRNAs. The pre-miRNAs are exported to the cytoplasm by Exportin-5, where they are subsequently processed by Dicer into mature miRNAs

The pre-miRNAs which are exported to the cytoplasm, contain double stranded stem regions that are cleaved by Dicer to generate mature miRNAs (Fig. 3).

Drosha does not function in isolation *in vivo*, but rather is part of a multiprotein complex that has been termed Microprocessor (Denli *et al*., 2004). The Microprocessor complex, in addition to containing Drosha, includes a dsRNA-binding protein Pasha (partner of Drosha). Pasha is important for efficient processing of miRNA precursors as evidenced by the observation that biochemical or genetic depletion of Pasha activity, interferes with pri-miRNA processing (Denli *et al*., 2004; Han *et al.*, 2004; Landthaler *et al*., 2004).

2.4. Dicer

Combined biochemical and candidate gene selection approaches were used to identify Dicer as the bidentate ribonuclease activity that generates 22 nt siRNAs from long dsRNA precursors (Bernstein *et al*., 2001). Dicer also functions downstream of Drosha (Fig. 3) to produce mature miRNAs which like siRNAs, are incorporated into gene-silencing complexes called *R*NA-*i*nduced *s*ilencing *c*omplexes (RISCs) (Bernstein *et al*., 2001; Grishok *et al*., 2001; Hutvagner *et al*., 2001; Ketting *et al*., 2001).

Dicer enzymes are modular proteins that are composed of up to five different domains (Fig. 1). Of critical importance to Dicer function is the PAZ domain, a 130 amino acid residue domain located in the center of the Dicer sequence. This domain is also present in members of the Argonaute superfamily of proteins (Cerutti *et al.*, 2000). Together, the Dicer and Argonaute family members comprise the core components of the RNAi machinery. Originally thought to mediate protein-protein interactions between Dicer and Argonaute proteins, structural studies revealed that PAZ domains are in fact oligonucleotide-binding modules (Lingel *et al*., 2003; Song *et al*., 2003; Yan *et al*., 2003). Dicer appears to have specific requirements for substrate binding and there is considerable evidence indicating that PAZ domains are involved in binding the ends of small dsRNA molecules (Lingel *et al*., 2003; Song *et al*., 2003; Yan *et al*., 2003). In this respect, the characteristic 5' phosphates and 2 nt 3' overhangs that result from Drosha-mediated cleavage of pri-miRNAs,

are important for recognition by Dicer (Lee *et al*., 2003; Zhang *et al*., 2004). Through binding to the ends of dsRNA molecules, PAZ domains help to position the substrates for cleavage by the RNase III domains (Fig. 4). It is interesting to note that Dicer PAZ domains contain a large extended loop that is absent from Argonaute PAZ domains (MacRae *et al*., 2006). This difference between Dicer and Argonaute PAZ domains may be an important for differential binding and potentially transferring RNA substrates between these proteins.

The amino-terminal regions of most Dicer enzymes contain a DExH-box RNA helicase domain. DExH type helicases contain eight conserved motifs, and are named so because of the conservation of the amino acid sequence Asp-Glu-X-His in the fourth motif (Gorbalenya *et al*., 1989; Linder *et al*., 1989). Interestingly, while the sequences of the helicase domains are highly conserved in Dicer family members (Nicholson and Nicholson, 2002), ATP is not required for the RNase activity of recombinant mammalian Dicer *in vitro* (Provost *et al*., 2002; Zhang *et al.*, 2002). In fact, some Dicer enzymes in lower eukaryotes lack helicase domains altogether, yet these enzymes are fully capable of "dicing" dsRNA into uniformly sized products (MacRae *et al*., 2006).

While Dicer cleavage products are dsRNA molecules, one strand of the duplex, the "guide strand" is used to target RISCs to homologous mRNAs (Martinez *et al.*, 2002). The remaining strand, the "passenger strand", does not function in

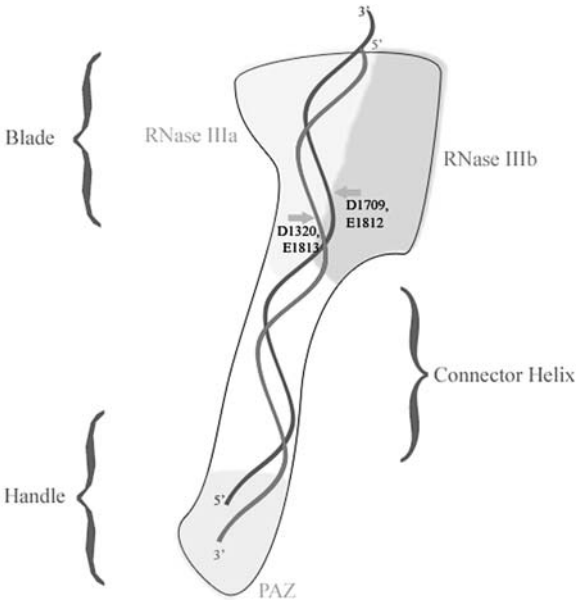


Figure 4. Model of Dicer cleavage. *Giardia* Dicer forms a hatchet-like structure with the two RNase III domains forming the blade and the PAZ domain forming the handle. One end of the dsRNA substrate is held by the PAZ domain and positioned approxiamtely 20–25 nt away from the active center in the blade region. The catalytic amino acids in the processing center are indicated

RISC targeting. A long-standing issue in RNAi biology has been to understand how unwinding of the siRNA duplex occurs following Dicer cleavage. Up until very recently, the prevailing theory held that ATP-dependent helicase activity was required to unwind siRNA duplexes after which the guide strand was loaded onto RISC (Bartel, 2004; Meister and Tuschl, 2004; Sontheimer, 2005; Tomari and Zamore, 2005). Certainly, the helicase of Dicer was considered as a likely entity to perform this function. Studies in *Drosophila melanogaster* revealed that the RNA helicase activity of Dicer-2 is necessary for siRNA production *in vivo*, however, it is not required for the downstream function of Dicer-2 (Lee *et al.*, 2003). In mammalian cells, incorporation of miRNA guide strands into RISC was found to occur in the absence of ATP (Maniataki and Mourelatos, 2005). Indeed, current evidence suggests that the helicase domain of Dicer is not required for unwinding dsRNAs before loading onto RISCs. Rather it seems that both the guide strand and passenger strands of Dicer cleavage products are loaded onto RISC. The passenger strand is then subsequently cleaved by the RISC leaving the guide strand available for targeting the mature complex to homologous mRNAs (Leuschner *et al.*, 2006; Matranga *et al.*, 2005; Rand *et al.*, 2005).

The DUF283 domain is approximately 100 amino acid residues in length and is also very well conserved in Dicer family members. It is located downstream of the helicase domain in the primary sequence of Dicer. The structure of this domain has not been solved yet and as the name implies, its role in Dicer function is not known.

The RNase III domains form the catalytic centers of Dicer enzymes. In all RNase III family members, dimerization of these domains is required for RNase activity (Dunn, 1976; Robertson *et al.*, 1968). Dimerization results in alignment of the active sites from each RNase III domain thereby creating a catalytic center. Such an arrangement allows these enzymes to cleave two nearby phosphodiester bonds on opposite strands of the duplex. This type of cleavage gives rise to products with 2 nt 3' protruding ends (Fig. 2). Based on the crystal structure of *Aquifex aeolicus* RNase III, Blaszczyk *et al.* suggested that the intersubunit cleft between two RNase III molecules, contains two compound catalytic centers positioned at the ends of the cleft. Two clusters of acid amino acid residues within the catalytic centers are in turn required for cleavage of the phosphodiester bonds (E40, D44, D107 and E110) and coordinating single Mg^{2+} or Mn^{2+} ions (E37 and E64). While spatially conserved acidic amino acid residues corresponding to E37, D44, D107 and E110 are present in the RNase IIIa domain of Dicer enzymes, amino acid residues equivalent to E37 and E64 are not conserved in RNase IIIb. These observations led to the idea that only one of the two "active" sites in the RNase IIIb domain is functional (Blaszczyk *et al.*, 2001; Hannon, 2002; Nicholson, 2003; Zamore, 2001).

The carboxyl termini of most Dicer enzymes include a double strand RNA-binding domain known as a dsRBD. These domains are relatively small in size (100 amino acid residues) and are conserved among eukaryotes and prokaryotes (Fierro-Monti and Mathews, 2000; Green and Mathews, 1992; St Johnston *et al.*, 1992). The structures of dsRBDs have been studied by X-ray crystallography and by nuclear magnetic resonance. They adopt an $\alpha\beta\beta\beta\alpha$ fold, a structure that is common

among oligonucleotide binding pockets (Carlson *et al*., 2003; Saunders, 2003). The protein/RNA interface that results from contact between a dsRBD and a dsRNA, spans two minor and one major groove of the RNA helix (Ryter and Schultz, 1998). It is likely that dsRBDs differentiate dsRNA from DNA by recognition of the 2' hydroxyl groups on RNA molecules. Specifically, the amino terminal helix of the dsRBD binds the 2' hydroxyl groups that line the minor groove of dsRNA (Bevilacqua and Cech, 1996; Ryter and Schultz, 1998). In addition, recent studies have revealed that some dsRBDs contain a carboxyl terminal helix that binds to hairpin structures (Leulliot *et al*., 2004). In theory, the dsRBD represents a means whereby Dicer binds indiscriminately to dsRNA substrates, however, dsRBDs are not strictly required for the function of Dicer *in vivo*. For example, the *Giardia intestinalis* Dicer homolog does not contain a dsRBD, but is capable of "dicing" dsRNA substrates into 25 nt products (MacRae *et al*., 2006). Accordingly, the role, if any, of this domain in selectively retaining the dsRNA substrates for Dicer cleavage (*i.e.* dsRNA containing 2 nt 3' overhangs, 5' phosphate groups and 3' hydroxyl groups) is unknown.

2.5. Mechanism of Dicer Cleavage

As stated above, the RNase III domains of Dicer and other RNase III superfamily enzymes are only active when they are in dimeric form. Unlike bacterial RNase III enzymes whose cleavage products are typically 11 nt in length, Dicer cleavage products are 21–25 nt. Accordingly, models of Dicer cleavage must account for the differences in product lengths. Two early models proposed that Dicer functions as a homodimer that is organized in antiparallel or head to tail orientations (reviewed in Carmell and Hannon, 2004). A major concern regarding these models is that they both require tight packing of the two Dicer molecules along the dsRNA substrate. Because of the potential for steric hindrance between the RNase III domains and the large amino-terminal region that contains the helicase, DUF283 and PAZ domains, it is uncertain as to whether the RNase III domains of two Dicer molecules can fit into a small enough space to generate 22 nt products. In fact, if the steric relationship between the active domains is similar to the *E. coli* RNase III active dimer, the products of cleavage are expected to be approximately 30 nt long, rather than 21–25 nt (Carmell and Hannon, 2004). Finally, and perhaps more importantly, current evidence suggests that Dicer exists as a monomer (Zhang *et al*., 2004).

Although the structure of a mammalian or invertebrate Dicer molecule has yet to be solved, MacRae *et al*. (2006) recently determined the 3.3 Å resolution structure of a Dicer homolog from *Giardia*. Interestingly, this enzyme is considerably smaller than most Dicer proteins and lacks helicase, DUF 283 and dsRBD domains. However, it is highly active and is able to generate 25 nt products from long dsRNA substrates. The results obtained from the functional analyses of the *Giardia* Dicer protein, are consistent with data from the Filipowicz group that support the idea of a single processing center model (Zhang *et al*., 2004). Central to this model is the observation that Dicer cleavage occurs from the ends of dsRNA molecules

(Zhang *et al*, 2002). The *Giardia* Dicer is an elongated molecule that resembles a hatchet (Fig. 4). The two RNase III domains are arranged to form an intramolecular dimer with a single processing center resembling a "blade" (Fig. 4). The RNase blade region is linked to the PAZ-containing handle region by a connector helix. The distance from the RNase III blade region and the PAZ domain is thought to govern the length of the RNA cleavage products. In this respect, Dicer has been compared to a molecular ruler (Carmell and Hannon, 2004; MacRae *et al*, 2006; Zhang *et al*, 2004). In the *Giardia* Dicer molecule, the active site of the RNase III dimer is positioned 65 Å or approximately 25 nt away from the substrate terminus which is anchored by the PAZ domain (Fig. 4). Presumably, in mammalian Dicer molecules, the distance between the PAZ domain and RNase III processing center is shorter.

One potential flaw in this model is that not all Dicer molecules appear to contain PAZ domains. For example, Dicer from the fission yeast *Schizosaccharomyces pombe* does not contain a recognized PAZ domain, yet it is capable of generating siRNAs of the expected size (Reinhart and Bartel, 2002). However, due to the relatively poor primary sequence conservation among PAZ domains it is possible that the *S. pombe* Dicer contains a divergent form of the PAZ domain that is simply not recognized by sequence analyses algorithms.

3. DICER-BINDING PROTEINS

Given the enormous amount of interest in RNAi mechanisms, it is perhaps surprising that comprehensive genetic and proteomic screens for Dicer-interacting proteins in mammalian systems have not been reported yet. Recently, however, Mello and colleagues used a whole organism approach for identification of Dicer-binding proteins in *C. elegans* (Duchaine *et al*, 2006). The DCR-1 enzyme appears to interact with at least 20 different proteins. As expected, many of these proteins are known to be involved in RNAi pathways and include members of the Argonaute family, RNA helicases, dsRNA binding proteins and RNA-directed RNA polymerase. Three of the newly discovered DCR-1 binding proteins were studied in more detail and of these, the PIR-1 phosphatase is perhaps most intriguing because up to this point, not much thought had been given to the potential roles of kinases and phosphatases in RNAi. PIR-1 is not required for processing of the trigger dsRNAs by DCR-1, but amplified dsRNA intermediates are not cleaved by DCR-1 unless PIR-1 is present. Based on how mammalian homologs of PIR-1 function (Deshpande *et al*, 1999; Yuan *et al*, 1998), it is thought that PIR-1 processes the termini of amplified dsRNAs to generate optimal 5' ends that are recognized by DCR-1 and possibly Argonaute proteins.

3.1. RISC

Argonaute proteins are the most well characterized group of Dicer-binding proteins. Togther with Dicer, they form the core components of RISCs, which are the effectors

of gene-silencing. Biochemical and genetic assays revealed that Dicer/Argonaute interactions are mediated by the RNase III and PIWI domains of Dicer and Argonaute respectively (Tahbaz *et al*., 2004). Interestingly, binding of Argonaute proteins inhibits the RNase activity of Dicer *in vitro*. One interpretation of these data, is that the interaction serves to maintain the specificity of a given RISC for a single mRNA species by preventing Dicer from engaging other RNA substates. Activity of the heat-shock protein 90 (Hsp90) was found to be required for stable binding of Dicer to Argonaute (Tahbaz *et al*., 2004). The involvement of Hsp90 in assembly of RISC components suggests that formation of gene-silencing complexes is highly regulated.

Loading of the guide strand onto Argonaute proteins is required for targeting mature RISCs to specific mRNAs. This process is discussed in further detail in the next section. The fate of the RISC-targeted mRNA depends on two major factors. First, in order for RNA-directed cleavage to occur, siRNAs must be bound to a cleavage competent Argonaute protein. Of the four mammalian Argonaute proteins, only Ago2 exhibits cleavage activity (Liu *et al*., 2004). The second factor is the level of sequence identity between the guide strand RNA and the mRNA (Elbashir *et al.*, 2001a; Elbashir *et al.*, 2001b; Paddison *et al*., 2002). If the guide strand is 100% identical to the mRNA and base pairing is complete, the mRNA will be cleaved by the endonuclease activity of Ago2 (Fig. 5). If, however, the base pairing is incomplete, and the RISC binds to the 3' untranslated region of the mRNA, translation is prevented without concomitant degradation (Elbashir *et al.*, 2001a; Elbashir *et al.*, 2001b; Paddison *et al*., 2002). Recent evidence suggests that in this context, Argonaute proteins prevent translation by interfering with recognition of the cap structure on the mRNA (Pillai *et al*., 2005).

3.2. The RISC Loading Complex (RLC)

Dicer cleavage products (siRNAs and miRNAs) require the action of a multi-protein complex, the *R*ISC *l*oading *c*omplex (RLC) for functional interaction with Argonaute proteins. The RLC is best studied in *D. melanogaster*, whose genome encodes two Dicer paralogs. Dicer-2 is required for siRNA-mediated RISC activity, while Dicer-1 is required for miRNA-mediated RISC activity (Lee *et al*., 2004). The *D. melanogaster* RLC includes Dicer-2 and its cognate binding protein R2D2 (Liu *et al*., 2004). R2D2 is a RNA-binding protein that contains tandem dsRBDs. It is required for sensing the thermodynamic asymmetry of RNA duplexes (Tomari *et al.*, 2004a; Tomari *et al*., 2004b). The ends of asymmetric RNA duplexes have inherently different thermodynamic properties, and it is hypothesized that the guide strand is chosen based on these differences. Within the RLC, R2D2 is thought to bind to the more thermodynamically stable end of the siRNA duplex, and in doing so, displaces Dicer-2, thereby ensuring that it binds to the less thermodynamically stable end of the duplex.

Loquacious is a dsRBD-containing protein that binds to *D. melanogaster* Dicer-1. The functional relevance of Loquacious is illustrated by the observation that

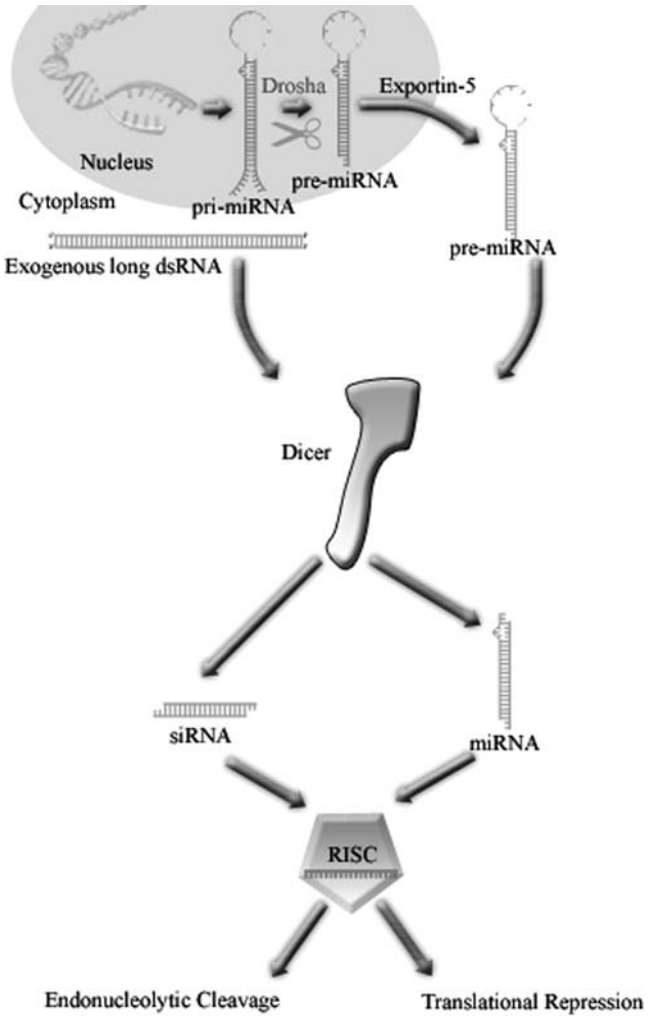


Figure 5. Overview of RNA-dependent post-transcriptional gene-silencing. Drosha generated pre-miRNAs or long dsRNA are processed by Dicer in the cytoplasm to produce miRNAs or siRNAs respectively. The guide strands of the Dicer products target RISCs to homologous mRNAs that are either cleaved or translationally repressed by the Argonaute subunit

Dicer-1-dependent maturation of pre-miRNAs requires interaction with this protein (Forstemann *et al*., 2005; Saito *et al*., 2005). Not surprisingly, binding partners of mammalian Dicer enzymes, are now being identified. For example, the Loquacious homolog, trans-activating response RNA-binding protein (TRBP) plays a role in siRNA- and miRNA- mediated gene silencing (Chendrimada *et al*., 2005; Haase *et al.*, 2005). In this case, interaction between TRBP and mammalian Dicer appears

to be required for the assembly of the RISC complex (Chendrimada *et al*, 2005). In addition, PACT, a multiple dsRBD-containing protein, binds both Dicer and Ago2 in a 500 kDa protein complex (Lee *et al*, 2006). Interactions between the helicase-containing region of Dicer and the most carboxyl terminal dsRBDs of PACT and TRBP are required for complex formation. Because TRBP and PACT bind to the same region of Dicer, it has been proposed that RISC assembly in mammalian cells is promoted by different Dicer-binding proteins (Lee *et al*, 2006).

Instead of helicase-mediated unwinding of dsRNA before loading onto RISC, it is now believed that RISC maturation requires Argonaute-mediated cleavage of the passenger strand (Matranga *et al*, 2005; Rand *et al*, 2005). Together, these studies indicate that passenger strand cleavage plays a significant role in the formation of mature RISC; however, the existence of a slower, cleavage-independent "bypass" mechanism that facilitates RISC assembly has also been proposed (Matranga *et al*, 2005).

4. DICER AND THE APPLICATION OF GENE-SILENCING METHODS

Shortly after the discovery of RNAi as a naturally occurring phenomenon, scientists soon realized that it could be used as an efficient gene-silencing tool in the laboratory. During this period of time, the development of RNAi, as a technique, rapidly outpaced the understanding of RNAi biology. For example, the introduction of *in vitro* transcribed long dsRNAs (ranging from approximately 500 to 3000 bp) into invertebrates by microinjection, was found to efficiently suppress the expression of homologous genes in invertebrates (Dzitoyeva *et al*, 2001). Later, it was discovered that simply "soaking" organisms such as nematodes in dsRNA-containing solutions or allowing them to feed on bacteria transformed with plasmids that encode long dsRNAs, resulted in robust and specific RNAi (Timmons and Fire, 1998). All of these methods of course, required the *in vivo* action of Dicer to cleave the long dsRNAs into siRNAs. By exploiting the RNAi pathway, it is theoretically possible to "knock down" the expression of any known gene. Indeed, the application of RNAi has been hailed as the new somatic cell genetics. As our understanding of gene-silencing pathways advanced, there has been a coincidental evolution of RNAi reagents that are used to "knock down" expression of genes. Below, we discuss the evolution of RNAi technology for use in post-transcriptional gene-silencing. For the sake of brevity, much of the discussion is focused on the use of RNAi in mammalian cells.

4.1. Long dsRNAs

The first RNAi reagents were simply long dsRNA molecules, which, when introduced into cells or organisms, resulted in silencing of homologous mRNAs. Wide-spread use of these reagents followed the discovery that dsRNA molecules were the triggers for RNAi. In this seminal work, Fire and colleagues observed

robust and specific silencing of the *unc-22* gene after they introduced a 742 base pair segment of homologous RNA into *Caenorhabditis elegans* (Fire *et al*, 1998). Because of their lengths, it was not economically feasible to chemically synthesize these reagents for routine laboratory use. However, the dsRNAs can be easily produced using T7 RNA polymerase in *in vitro* transcription reactions. Long dsRNAs can be introduced into cells or organisms by microinjection or transfection with cationic lipids. In addition, RNAi can be induced in nematodes through ingestion of bacteria that express long dsRNAs (Timmons and Fire, 1998). These reagents work very well for gene-silencing in well-characterized invertebrate systems such as *C. elegans* and *D. melanogaster*. However, there are two major limitations with the use of long dsRNAs. First, the RNA-dependent silencing is inherently transient. Unless the target cell has a means to amplify the siRNAs (*e.g.* RNA dependent RNA polymerase), they will be degraded or their effects lost as a result of dilution after cell division. Second, this method is not suitable for use in most mammalian cells because they activate the innate immune system. Perhaps because the replication intermediates of many RNA viruses are double stranded RNAs, the appearance of long dsRNAs (over 30 bp) in mammalian cells can trigger a double-stranded RNA-dependent protein kinase (PKR)-dependent anti-viral response. In this case, total cellular translation is inhibited following phosphorylation of eIF2α (Manche *et al*, 1992; Minks *et al*, 1979). The exceptions are cell lines that lack the ability to mount PKR or interferon responses (*e.g.* embryonic cell lines).

4.2. Synthetic siRNAs

A major advance in the use of dsRNA for gene-silencing followed the realization that Dicer cleavage products (siRNAs) are important effectors of RNAi (Bernstein *et al.*, 2001; Hammond *et al*, 2000). This led to the wide-spread use of chemically synthesized siRNAs as silencing triggers. The rationale for this approach was that synthetic siRNAs resemble Dicer cleavage products and that their incorporation into RISC particles would direct the accurate targeting and destruction of homologous mRNAs. The procedure involves transfection of 21–22 nt synthetic siRNAs into target cells. Typically, the siRNAs share 100% sequence identity to a region of the targeted mRNA, often within the coding region. These dsRNAs are synthesized as two separate 21-mers that share 19 bp of sequence complementarity and form 19 bp of dsRNA with 2 nt 3' overhangs. The separate 21-mers are then annealed before transfection. Because of their length and the characteristic 3' overhangs, the annealed 21-mers mimic Dicer cleavage products and provide robust and specific gene-silencing (Elbashir *et al.*, 2001a). Since they are under 30 base pairs, transfected siRNAs, in most cases, do not induce a PKR-dependent global shutdown of cellular translation. The siRNAs can be purchased from commercial suppliers at a reasonable cost and are very easy to use. As a result, many laboratories with no prior knowledge of RNAi biology have been able to exploit RNAi as a gene-silencing tool.

There are, however, two significant drawbacks to this method. First, unlike nematodes and plants, mammalian cells do not have a mechanism to amplify the siRNA triggers and, therefore, sustained gene-silencing can only be achieved through repeated transfection with siRNA. Second, different siRNAs directed against the same mRNA often exhibit different silencing efficiencies. This problem has been alleviated in large part by the use of algorithms for design and selection of siRNAs. A number of these programs are freely available online. One need only enter the nucleotide sequence or genebank accession number of the mRNA into the algorithm and a number of candidate siRNA sequences are provided.

Based on the canonical RNAi pathway, 21–22 nt siRNAs are not expected to require the action of Dicer to effect gene-silencing. However, in addition to its well recognized role in initiation of RNAi, evidence points to a downstream role for Dicer in RISC function (Pham *et al*., 2004). Indeed, down-regulation of Dicer appears to decrease the efficiency of siRNA-mediated gene-silencing in mammalian cells (Doi *et al.*, 2003). It is also interesting to note that slightly longer siRNAs (25–30 nt) are reportedly up to 100-fold more potent than 21-mers of similar sequence (Kim *et al*., 2005). Given that Dicer and Argonaute proteins interact in the context of a RISC, it was hypothesized that the longer siRNAs, which are substrates for Dicer, are better silencing reagents because the production of 21 nt siRNAs by Dicer is now directly coupled to their loading onto the Argonaute subunit of RISC. Thus, for a slight increase in the cost of the siRNAs, researchers can benefit from a significant increase in gene-silencing efficiency.

Finally, a third siRNA-based method involves the *in vitro* production of transfectable pools of siRNAs. Long dsRNAs, synthesized in a T7 polymerase-dependent reaction, are subjected to Dicer cleavage *in vitro* and the resulting pools of siRNAs are transfected into target cells (Kawasaki *et al*., 2003).Recombinant Dicer is available from a number of commercial suppliers. The major advantage of "*in vitro* dicing" is that unlike using pre-designed chemically synthesized siRNAs, this method does not require optimizing target sites within mRNAs.

4.3. Small Hairpin RNA (shRNA)

The third generation of RNAi-based gene-silencing reagents is based on the use of plasmid or viral vectors for the production of Dicer substrates. Because selectable markers can be included in the vectors, these reagents allow for prolonged gene-silencing through the production of stable cell lines that express the Dicer substrates. The first plasmid-based silencing systems typically employed RNA pol III promoters for production of short hairpin RNAs (shRNAs). The shRNAs are designed to encode inverted repeats of 19–29 bp separated by a 6–9nt loop. The resulting transcripts fold into hairpin structures that resemble pre-miRNAs, which are Drosha cleavage products (Fig. 3). Accordingly, these hairpins are recognized as substrates for Dicer and concomitant cleavage by this enzyme yields 21–25 nt products that are incorporated into RISCs. Alternatives to this scheme include co-transfection of two plasmids that each encode one strand of a RNA duplex.

Recently, a new generation of shRNA-based vectors has been developed and these reagents appear to be significantly more potent than first generation shRNA reagents (Dickins *et al*, 2005; Silva *et al*, 2005). These new vectors encode shRNAs that resemble pri-miRNAs, as opposed to the first generation vectors that encode pre-miRNA-like transcripts. Therefore rather than serving directly as Dicer substrates, the pri-miRNA-like silencing triggers are first substrates for Drosha, followed by a Dicer-dependent processing step. Through coupling of the Drosha and Dicer processing steps, the pri-miRNA-like shRNAs are reportedly up to 12 times more potent than comparable first generation shRNAs (Silva *et al*, 2005). Making life even easier for scientists is the fact that a number of libraries containing pre-designed shRNA constructs, which cover most of the human and mouse genomes, are now available (Silva *et al*, 2005). Finally, because the pri-miRNA-like shRNAs can be expressed from RNA pol II promoters, RNAi-dependent gene-silencing can now be tightly regulated. For example, tetracycline-responsive promoters have been used to drive the expression of pri-miRNA-like shRNAs in both cultured cells and mice (Dickins *et al*, 2005). The creation of inducible RNAi tools adds a whole new dimension to this technology. Specifically, by introducing tissue-specific promoters, one can limit knock down of genes to tissues of interest. In addition, by fine tuning the level of knock down, it should be easier to study genes that are critical for viability.

ACKNOWLEDGEMENTS

This work was funded by grants from the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council and the Cancer Research Society. T.C.H. is the recipient of a Medical Scientist award from the Alberta Heritage Foundation for Medical Research.

REFERENCES

Bartel, D.P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. *116*, 281–97.

Bernstein, E., A.A. Caudy, S.M. Hammond and G.J. Hannon. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. *409*, 363–6.

Bevilacqua, P.C. and T.R. Cech. 1996. Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. Biochemistry. *35*, 9983–94.

Blaszczyk, J., J.E. Tropea, M. Bubunenko, K.M. Routzahn, D.S. Waugh, D.L. Court and X. Ji. 2001. Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. Structure. *9*, 1225–36.

Bohnsack, M.T., K. Czaplinski and D. Gorlich. 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. Rna. *10*, 185–91.

Bram, R.J., R.A. Young and J.A. Steitz. 1980. The ribonuclease III site flanking 23S sequences in the 30S ribosomal precursor RNA of E. coli. Cell. *19*, 393–401.

Carlson, C.B., O.M. Stephens and P.A. Beal. 2003. Recognition of double-stranded RNA by proteins and small molecules. Biopolymers. *70*, 86–102.

Carmell, M.A. and G.J. Hannon. 2004. RNase III enzymes and the initiation of gene silencing. Nat Struct Mol Biol. *11*, 214–8.

Catala, M., B. Lamontagne, S. Larose, G. Ghazal and S.A. Elela. 2004. Cell cycle-dependent nuclear localization of yeast RNase III is required for efficient cell division. Mol Biol Cell. *15*, 3015–30.

Cerutti, L., N. Mian and A. Bateman. 2000. Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain. Trends Biochem Sci. *25*, 481–2.

Chendrimada, T.P., R.I. Gregory, E. Kumaraswamy, J. Norman, N. Cooch, K. Nishikura and R. Shiekhattar. 2005. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature. *436*, 740–4.

Dasgupta, S., L. Fernandez, L. Kameyama, T. Inada, Y. Nakamura, A. Pappas and D.L. Court. 1998. Genetic uncoupling of the dsRNA-binding and RNA cleavage activities of the Escherichia coli endoribonuclease RNase III–the effect of dsRNA binding on gene expression. Mol Microbiol. *28*, 629–40.

Denli, A.M., B.B. Tops, R.H. Plasterk, R.F. Ketting and G.J. Hannon. 2004. Processing of primary microRNAs by the Microprocessor complex. Nature. *432*, 231–5.

Deshpande, T., T. Takagi, L. Hao, S. Buratowski and H. Charbonneau. 1999. Human PIR1 of the protein-tyrosine phosphatase superfamily has RNA 5'-triphosphatase and diphosphatase activities. J Biol Chem. *274*, 16590–4.

Dickins, R.A., M.T. Hemann, J.T. Zilfou, D.R. Simpson, I. Ibarra, G.J. Hannon and S.W. Lowe. 2005. Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. Nat Genet. *37*, 1289–95.

Doi, N., S. Zenno, R. Ueda, H. Ohki-Hamazaki, K. Ui-Tei and K. Saigo. 2003. Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. Curr Biol. *13*, 41–6.

Duchaine, T.F., J.A. Wohlschlegel, S. Kennedy, Y. Bei, D. Conte, Jr., K. Pang, D.R. Brownell, S. Harding, S. Mitani, G. Ruvkun, J.R. Yates, 3rd and C.C. Mello. 2006. Functional Proteomics Reveals the Biochemical Niche of C. elegans DCR-1 in Multiple Small-RNA-Mediated Pathways. Cell. *124*, 343–54.

Dunn, J.J. 1976. RNase III cleavage of single-stranded RNA. Effect of ionic strength on the fideltiy of cleavage. J Biol Chem. *251*, 3807–14.

Dzitoyeva, S., N. Dimitrijevic and H. Manev. 2001. Intra-abdominal injection of double-stranded RNA into anesthetized adult Drosophila triggers RNA interference in the central nervous system. Mol Psychiatry. *6*, 665–70.

Elbashir, S.M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber and T. Tuschl. 2001a. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature. *411*, 494–8.

Elbashir, S.M., J. Martinez, A. Patkaniowska, W. Lendeckel and T. Tuschl. 2001b. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. Embo J. *20*, 6877–88.

Fierro-Monti, I. and M.B. Mathews. 2000. Proteins binding to duplexed RNA: one motif, multiple functions. Trends Biochem Sci. *25*, 241–6.

Fire, A., S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver and C.C. Mello. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. *391*, 806–11.

Forstemann, K., Y. Tomari, T. Du, V.V. Vagin, A.M. Denli, D.P. Bratu, C. Klattenhoff, W.E. Theurkauf and P.D. Zamore. 2005. Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. PLoS Biol. *3*, e236.

Fortin, K.R., R.H. Nicholson and A.W. Nicholson. 2002. Mouse ribonuclease III. cDNA structure, expression analysis, and chromosomal location. BMC Genomics. *3*, 26.

Gegenheimer, P. and D. Apirion. 1980. Precursors to 16S and 23S ribosomal RNA from a ribonuclear III-strain of Escherichia coli contain intact RNase III processing sites. Nucleic Acids Res. *8*, 1873–91.

Gegenheimer, P. and D. Apirion. 1981. Processing of procaryotic ribonucleic acid. Microbiol Rev. *45*, 502–41.

Gerbi, S. 1995. Expansion segments: regions of variable size that interrupt the universal core secondary structure of ribosomal RNA. *In* Ribosomal RNA - Structure, Evolution, Processing, and Function in Protein Biogenesis. A.D. RA Zimmermann, editor. CRC Press, Boca Raton, New York. 71–87.

Gorbalenya, A.E., E.V. Koonin, A.P. Donchenko and V.M. Blinov. 1989. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res. *17*, 4713–30.

Gray, M.a.M.S. 1995. Evolution of rRNA gene organization. *In* Ribosomal RNA - Structure, Evolution, Processing, and Function in Protein Biogenesis. A.D. RA Zimmermann, editor. CRC Press, Boca Raton, New York. 46–69.

Green, S.R. and M.B. Mathews. 1992. Two RNA-binding motifs in the double-stranded RNA-activated protein kinase, DAI. Genes Dev. *6*, 2478–90.

Grishok, A., A.E. Pasquinelli, D. Conte, N. Li, S. Parrish, I. Ha, D.L. Baillie, A. Fire, G. Ruvkun and C.C. Mello. 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell. *106*, 23–34.

Haase, A.D., L. Jaskiewicz, H. Zhang, S. Laine, R. Sack, A. Gatignol and W. Filipowicz. 2005. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO Rep. *6*, 961–7.

Hammond, S.M., E. Bernstein, D. Beach and G.J. Hannon. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature. *404*, 293–6.

Han, J., Y. Lee, K.H. Yeom, Y.K. Kim, H. Jin and V.N. Kim. 2004. The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev. *18*, 3016–27.

Hannon, G.J. 2002. RNA interference. Nature. *418*, 244–51.

Hutvagner, G., J. McLachlan, A.E. Pasquinelli, E. Balint, T. Tuschl and P.D. Zamore. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science. *293*, 834–8.

Kawasaki, H., E. Suyama, M. Iyo and K. Taira. 2003. siRNAs generated by recombinant human Dicer induce specific and significant but target site-independent gene silencing in human cells. Nucleic Acids Res. *31*, 981–7.

Ketting, R.F., S.E. Fischer, E. Bernstein, T. Sijen, G.J. Hannon and R.H. Plasterk. 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. *15*, 2654–9.

Kim, D.H., M.A. Behlke, S.D. Rose, M.S. Chang, S. Choi and J.J. Rossi. 2005. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. Nat Biotechnol. *23*, 222–6.

Landthaler, M., A. Yalcin and T. Tuschl. 2004. The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. Curr Biol. *14*, 2162–7.

Lee, Y., C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Radmark, S. Kim and V.N. Kim. 2003. The nuclear RNase III Drosha initiates microRNA processing. Nature. *425*, 415–9.

Lee, Y., I. Hur, S.Y. Park, Y.K. Kim, M.R. Suh and V.N. Kim. 2006. The role of PACT in the RNA silencing pathway. Embo J.

Lee, Y.S., K. Nakahara, J.W. Pham, K. Kim, Z. He, E.J. Sontheimer and R.W. Carthew. 2004. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell. *117*, 69–81.

Leulliot, N., S. Quevillon-Cheruel, M. Graille, H. van Tilbeurgh, T.C. Leeper, K.S. Godin, T.E. Edwards, S.T. Sigurdsson, N. Rozenkrants, R.J. Nagel, M. Ares and G. Varani. 2004. A new alpha-helical extension promotes RNA binding by the dsRBD of Rnt1p RNAse III. Embo J. *23*, 2468–77.

Leuschner, P.J., S.L. Ameres, S. Kueng and J. Martinez. 2006. Cleavage of the siRNA passenger strand during RISC assembly in human cells. EMBO Rep.

Lewis, B.P., C.B. Burge and D.P. Bartel. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. *120*, 15–20.

Linder, P., P.F. Lasko, M. Ashburner, P. Leroy, P.J. Nielsen, K. Nishi, J. Schnier and P.P. Slonimski. 1989. Birth of the D-E-A-D box. Nature. *337*, 121–2.

Lingel, A., B. Simon, E. Izaurralde and M. Sattler. 2003. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. Nature. *426*, 465–9.

Liu, J., M.A. Carmell, F.V. Rivas, C.G. Marsden, J.M. Thomson, J.J. Song, S.M. Hammond, L. Joshua-Tor and G.J. Hannon. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. Science. *305*, 1437–41.

Lund, E., S. Guttinger, A. Calado, J.E. Dahlberg and U. Kutay. 2004. Nuclear export of microRNA precursors. Science. *303*, 95–8.

Macrae, I.J., K. Zhou, F. Li, A. Repic, A.N. Brooks, W.Z. Cande, P.D. Adams and J.A. Doudna. 2006. Structural basis for double-stranded RNA processing by Dicer. Science. *311*, 195–8.

Manche, L., S.R. Green, C. Schmedt and M.B. Mathews. 1992. Interactions between double-stranded RNA regulators and the protein kinase DAI. Mol Cell Biol. *12*, 5238–48.

Maniataki, E. and Z. Mourelatos. 2005. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. Genes Dev. *19*, 2979–90.

Martinez, J., A. Patkaniowska, H. Urlaub, R. Luhrmann and T. Tuschl. 2002. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell. *110*, 563–74.

Matranga, C., Y. Tomari, C. Shin, D.P. Bartel and P.D. Zamore. 2005. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell. *123*, 607–20.

Meister, G. and T. Tuschl. 2004. Mechanisms of gene silencing by double-stranded RNA. Nature. *431*, 343–9.

Minks, M.A., D.K. West, S. Benvin and C. Baglioni. 1979. Structural requirements of double-stranded RNA for the activation of 2',5'-oligo(A) polymerase and protein kinase of interferon-treated HeLa cells. J Biol Chem. *254*, 10180–3.

Nicholson, A.W. 1996. Structure, reactivity, and biology of double-stranded RNA. Prog Nucleic Acid Res Mol Biol. *52*, 1–65.

Nicholson, A.W. 2003. The ribonuclease III superfamily: forms and functions in RNA, maturation, decay, and gene silencing. *In* RNAi A Guide to Gene Silencing. G.J. Hannon, editor. Cold Spring Harbor Laboratory Press Cold Spring Harbor, NY. 149–174.

Nicholson, R.H. and A.W. Nicholson. 2002. Molecular characterization of a mouse cDNA encoding Dicer, a ribonuclease III ortholog involved in RNA interference. Mamm Genome. *13*, 67–73.

Paddison, P.J., A.A. Caudy, E. Bernstein, G.J. Hannon and D.S. Conklin. 2002. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. Genes Dev. *16*, 948–58.

Pham, J.W., J.L. Pellino, Y.S. Lee, R.W. Carthew and E.J. Sontheimer. 2004. A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. Cell. *117*, 83–94.

Pillai, R.S., S.N. Bhattacharyya, C.G. Artus, T. Zoller, N. Cougot, E. Basyuk, E. Bertrand and W. Filipowicz. 2005. Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science. *309*, 1573–6.

Provost, P., D. Dishart, J. Doucet, D. Frendewey, B. Samuelsson and O. Radmark. 2002. Ribonuclease activity and RNA binding of recombinant human Dicer. Embo J. *21*, 5864–74.

Rand, T.A., S. Petersen, F. Du and X. Wang. 2005. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. Cell. *123*, 621–9.

Regnier, P. and M. Grunberg-Manago. 1990. RNase III cleavages in non-coding leaders of Escherichia coli transcripts control mRNA stability and genetic expression. Biochimie. *72*, 825–34.

Reinhart, B.J. and D.P. Bartel. 2002. Small RNAs correspond to centromere heterochromatic repeats. Science. *297*, 1831.

Robertson, H.D., R.E. Webster and N.D. Zinder. 1968. Purification and properties of ribonuclease III from Escherichia coli. J Biol Chem. *243*, 82–91.

Ryter, J.M. and S.C. Schultz. 1998. Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. Embo J. *17*, 7505–13.

Saito, K., A. Ishizuka, H. Siomi and M.C. Siomi. 2005. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. PLoS Biol. *3*, e235.

Saunders, L.R. and G.N. Barber. 2003. The dsRNA binding protein family: critical roles, diverse cellular functions. Faseb J. *17*, 961–83.

Silva, J.M., M.Z. Li, K. Chang, W. Ge, M.C. Golding, R.J. Rickles, D. Siolas, G. Hu, P.J. Paddison, M.R. Schlabach, N. Sheth, J. Bradshaw, J. Burchard, A. Kulkarni, G. Cavet, R. Sachidanandam, W.R. McCombie, M.A. Cleary, S.J. Elledge and G.J. Hannon. 2005. Second-generation shRNA libraries covering the mouse and human genomes. Nat Genet. *37*, 1281–8.

Song, J.J., J. Liu, N.H. Tolia, J. Schneiderman, S.K. Smith, R.A. Martienssen, G.J. Hannon and L. Joshua-Tor. 2003. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nat Struct Biol. *10*, 1026–32.

Sontheimer, E.J. 2005. Assembly and function of RNA silencing complexes. Nat Rev Mol Cell Biol. *6*, 127–38.

St Johnston, D., N.H. Brown, J.G. Gall and M. Jantsch. 1992. A conserved double-stranded RNA-binding domain. Proc Natl Acad Sci U S A. *89*, 10979–83.

Steege, D.A. 2000. Emerging features of mRNA decay in bacteria. Rna. *6*, 1079–90.

Tahbaz, N., F.A. Kolb, H. Zhang, K. Jaronczyk, W. Filipowicz and T.C. Hobman. 2004. Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. EMBO Rep. *5*, 189–94.

Timmons, L. and A. Fire. 1998. Specific interference by ingested dsRNA. Nature. *395*, 854.

Tomari, Y., T. Du, B. Haley, D.S. Schwarz, R. Bennett, H.A. Cook, B.S. Koppetsch, W.E. Theurkauf and P.D. Zamore. 2004a. RISC assembly defects in the Drosophila RNAi mutant armitage. Cell. *116*, 831–41.

Tomari, Y., C. Matranga, B. Haley, N. Martinez and P.D. Zamore. 2004b. A protein sensor for siRNA asymmetry. Science. *306*, 1377–80.

Tomari, Y. and P.D. Zamore. 2005. Perspective: machines for RNAi. Genes Dev. *19*, 517–29.

Venema, J. and D. Tollervey. 1999. Ribosome synthesis in Saccharomyces cerevisiae. Annu Rev Genet. *33*, 261–311.

Wu, H., H. Xu, L.J. Miraglia and S.T. Crooke. 2000. Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. J Biol Chem. *275*, 36957–65.

Yan, K.S., S. Yan, A. Farooq, A. Han, L. Zeng and M.M. Zhou. 2003. Structure and conserved RNA binding of the PAZ domain. Nature. *426*, 468–74.

Yi, R., Y. Qin, I.G. Macara and B.R. Cullen. 2003. Exportin 5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev. *17*, 3011–6.

Yuan, Y., D.M. Li and H. Sun. 1998. PIR1, a novel phosphatase that exhibits high affinity to RNA . ribonucleoprotein complexes. J Biol Chem. *273*, 20347–53.

Zamore, P.D. 2001. Thirty-three years later, a glimpse at the ribonuclease III active site. Mol Cell. *8*, 1158–60.

Zhang, H., F.A. Kolb, V. Brondani, E. Billy and W. Filipowicz. 2002. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. Embo J. *21*, 5875–85.

Zhang, H., F.A. Kolb, L. Jaskiewicz, E. Westhof and W. Filipowicz. 2004. Single processing center models for human Dicer and bacterial RNase III. Cell. *118*, 57–68.

SECTION E

OXIDOREDUCTASES AND OTHER ENZYMES OF DIVERSE FUNCTION

CHAPTER 25

HYDROGEN PEROXIDE PRODUCING AND DECOMPOSING ENZYMES: THEIR USE IN BIOSENSORS AND OTHER APPLICATIONS

NÓRA ADÁNYI[1], TERÉZ BARNA[2], TAMÁS EMRI[3], MÁRTON MISKEI[3] AND ISTVÁN PÓCSI[3*]

[1] *Central Food Research Institute, Budapest,*
[2] *Department of Biochemistry and*
[3] *Department of Microbial Biotechnology and Cell Biology, Faculty of Science, University of Debrecen, Debrecen, Hungary*
[*] *istvanpocsi@yahoo.com*

1. INTRODUCTION

In this chapter H_2O_2-producing enzymes including glucose, galactose and cholesterol oxidases will be presented together with the H_2O_2 decomposer catalase. Industrial production of these redox enzymes is increasing rapidly owing to their established use in the construction of biosensors and novel applications, *e.g.* in the food and textile industries. We will cover three research areas to give a broad overview of these very important and remarkable enzymes, namely (i) studies on their structure and mechanism of action, (ii) the construction of oxidase- and catalase-based biosensors and (iii) novel applications.

2. ENZYME STRUCTURE AND MECHANISM OF ACTION

2.1. Glucose Oxidases

Glucose oxidase (β-D-glucose:oxygen 1-oxidoreductase, GOx, EC 1.1.3.4.) is a flavoprotein which catalyses the oxidation of β-D-glucose by O_2 to δ-gluconolactone and H_2O_2 (Fig. 1). The δ-gluconolactone subsequently hydrolyses to gluconic acid. Industrial production of GOx is mainly carried out using *Aspergillus niger* or *Penicillium amagasakiense* (Crueger and Crueger, 1990), however other *Penicillium* and *Aspergillus* spp. including *P. notatum, P. funiculosum, P. piceum,*

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 441–459

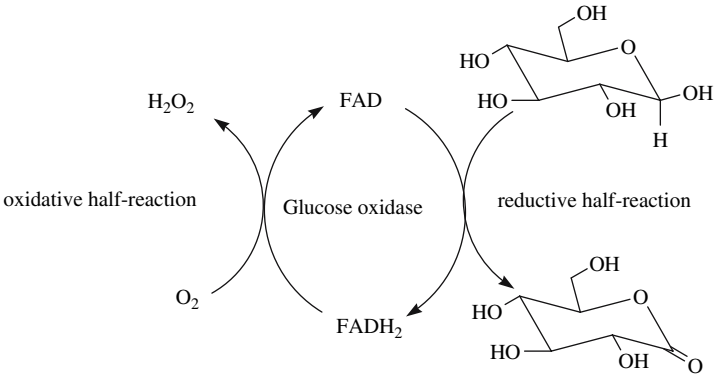


Figure 1. The reaction scheme for the oxidation of β-D-glucose catalysed by glucose oxidase

P. purpurogenum, P. variabile, P. chrysogenum and *A. fumaricus* are also good candidates for industrial applications (Crueger and Crueger, 1990; Petruccioli *et al*., 1993; Leiter *et al*., 2004). The *A. niger* and *P. amagasakiense* enzymes are homodimeric glycoproteins of 160 kDa molecular mass with one molecule of non-covalently bound FAD cofactor per subunit (Fig. 2; Wohlfahrt *et al*., 1999). Dissociation of subunits is only possible under denaturing conditions and is accompanied by the loss of the coenzyme FAD (Gouda *et al*., 2003). *A. niger* GOx has 10–16% mannose-type carbohydrate content, the carbohydrate moieties being *N*- or *O*-glycosidically linked to the protein (Takegawa *et al*., 1989). The two GOxs share 81% sequence similarity and exhibit similar glucose oxidation kinetics.

GOxs function by a Ping-Pong Bi-Bi mechanism which is divided into a reductive half-reaction and an oxidative half-reaction (Fig. 1). In the reductive half-reaction the substrate reduces FAD by hydride transfer forming dihydro-FAD ($FADH_2$) and the oxidized product. Numerous sugars and derivatives of β-D-glucose are the substrates of GOx but the enzyme shows far greater activity with β-D-glucose (Pazur and Kleppe, 1964). In the oxidative half reaction, the reduced flavin is regenerated by O_2 that undergoes reduction to H_2O_2. Besides O_2, both one-electron acceptors (such as transition metal complexes, ferrocene derivatives, and nitroxide radicals) and two-electron acceptors (benzoquinone derivatives) are good substrates for the reduced enzyme (Chan and Bruice, 1977; Bourdillon *et al*., 1993; Ryabov *et al*., 1999).

One-electron acceptors are generally used as mediators between GOx and the metal or graphite electrode. The largest mediator groups are derivatives of ferrocene that are small enough to penetrate to the active centre of GOx and exchange electrons (Alvarez-Icaza *et al*., 1995; Forrow *et al*., 2002). When the mediator cannot penetrate to the active site due to its bulkiness, a long-range electron transfer event takes place. Theoretical studies have predicted the length and path for electron transfer from the flavin cofactor to the surface of the protein (Alvarez-Icaza *et al*., 1995).

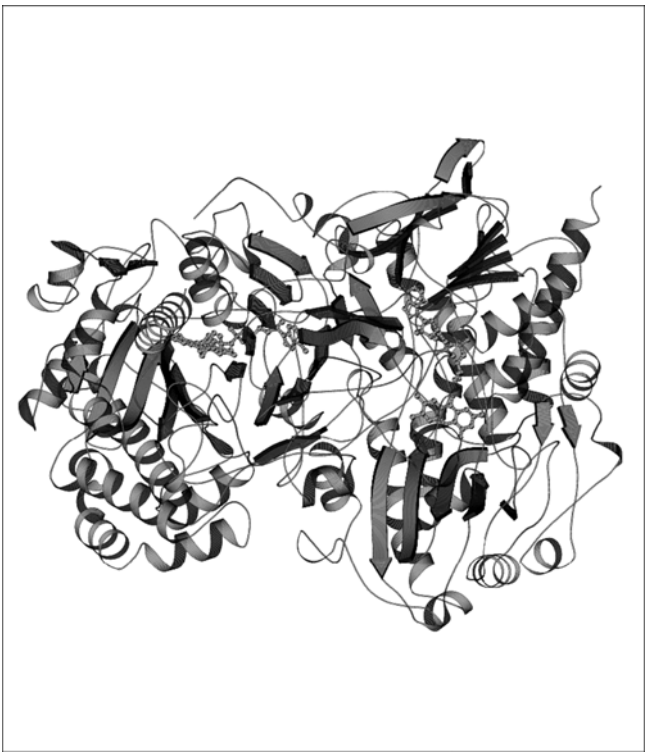


Figure 2. The homodimeric structure of glucose oxidase from *Penicillium amagasakiense* with non-covalently bound FAD represented by MOLSCRIPT (Kraulis, 1991)

2.2. Galactose Oxidases

Galactose oxidase (D-galactose:oxygen 6-oxidoreductase, GAO, EC 1.1.3.9.) is a Type II (non-blue) copper protein (Giordarno *et al*., 1974) with a molecular mass of 69 kDa. It catalyses the oxidation of a wide range of primary alcohols and polysaccharides (*e.g.* D-galactose) to the corresponding aldehydes, coupling this reaction to the reduction of O_2 to H_2O_2 : $RCH_2OH + O_2 \rightarrow RCHO + H_2O_2$ (Bretting and Jacobs, 1987). The substrate specificity of GAO is very broad, ranging from small alcohols to polysaccharides but the enzyme is highly stereospecific and does not oxidise either D-glucose or L-galactose (Klibanov *et al*., 1982; Goudsmit *et al*., 1984). The physiological function of GAO is unclear.

Different fungal species: *Fusarium dendroides (Dactylium dendroides), Gibberella fujikuroi* and *G. zeae* as well as the basidiomycete *Polyporus circinatus*, can be used for industrial-scale GAO production. GAO genes can also be expressed efficiently in either *Pichia pastoris* (Whittaker and Whittaker, 2000) or *Aspergillus oryzae* (Xu *et al*., 2000). GAO contains a single copper ion in addition to a protein base redox site (Whittaker and Whittaker, 1988, 1990) (Fig. 3), which comprises a

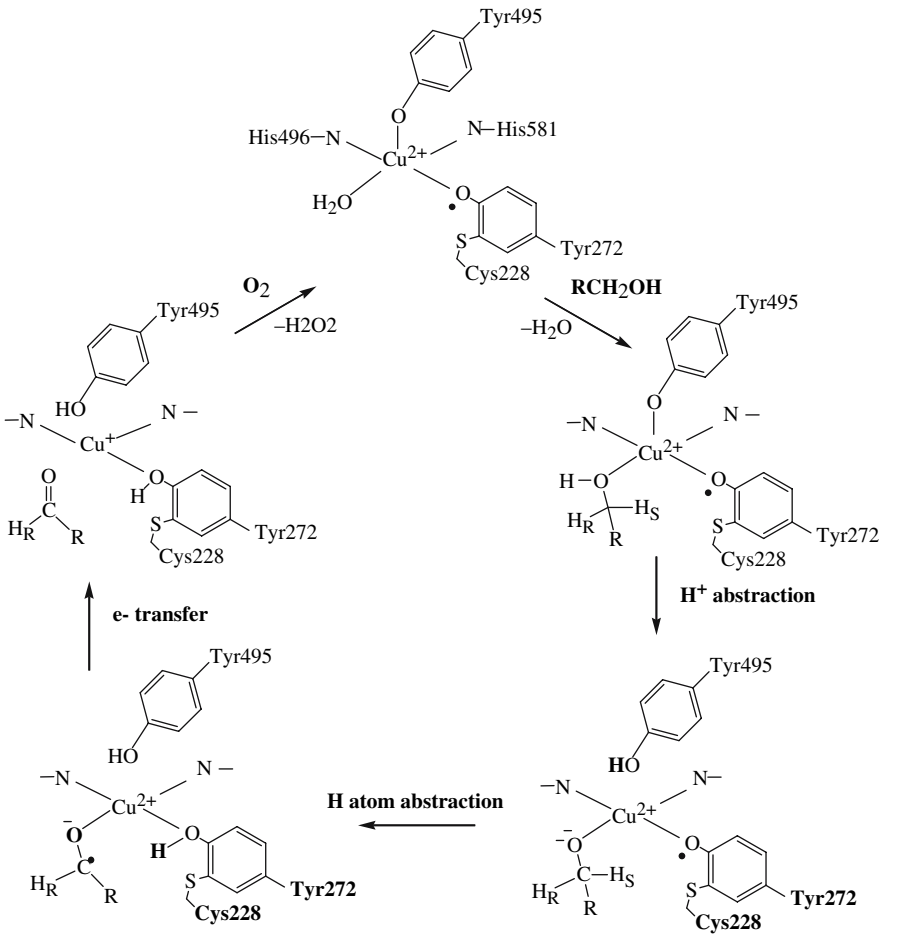


Figure 3. Proposed mechanism for D-galactose oxidase according to Whittaker (1988). R represents the pyranose ring associated with D-galactose

tyrosine residue (Tyr272) covalently linked to a cysteine (Cys228) through thioether bond, forming a cysteinyl-tyrosine (C-Y) "built-in" cofactor (Ito *et al*, 1991). The enzyme undergoes self-processing in the presence of oxygen and copper to generate this covalent cross-link, which is required for enzyme activity (Xie and van der Donk, 2001). The cross-linked tyrosine serves as an axial ligand to the copper and is oxidized to the radical form in the active state of the protein. The Cu(II)/Tyr cofactor on GAO carries out the two-electron oxidation of primary alcohols to the corresponding aldehydes *via* a radical mechanism.

The enzyme uses a Ping-Pong mechanism with respect to the substrate and O_2. Tyr495 functions as a base for abstracting a proton from the bound substrate followed by hydrogen atom abstraction by the Cys-Tyr radical generating an

alkoxide radical and diamagnetic Cys-Tyr (Whittaker and Whittaker, 1990). From the reactive alkoxide radical one electron transfers to Cu^{2+} reducing it to Cu^{+} and forming the aldehyde product. O_2 converts Cu^{+} back to Cu^{2+} (Fig. 3).

2.3. Cholesterol Oxidases

Cholesterol oxidases (cholesterol:oxygen oxidoreductase, ChOx, EC 1.1.3.6.) are bifunctional flavoenzymes that catalyse two reactions at a single active site. The first is the oxidation of cholesterol to cholest-5-ene-3-one and the isomerisation of the labile cholest-5-ene-3-one intermediate to cholest-4-ene-3-one product (Fig. 4). Although the enzymes exhibit a wide-range of steroid specificities, the 3β-hydroxyl group is essential for their activity (Inouye *et al*, 1982).

ChOxs are produced by bacteria including species from the *Rhodococcus, Streptomyces* and *Nocardia* genera. However, several *Brevibacterium, Proactinomyces, Pseudomonas* and *Cellulomonas* species also possess ChOx activity suitable for industrial production (Goodhue and Risley, 1978; Watanabe *et al*, 1986; MacLachlan *et al*, 2000). Recombinant *Escherichia coli* strains expressing bacterial ChOx genes have been reported (Sakka *et al*, 1994). Several fungi from the Basidiomycete taxon (*e.g. Lentinus edodes, Oudemansiella radicate, Coprinus comatus* and *Auricularia polytricha*) are also promising candidates for future industrial ChOx production (Matsui *et al*, 1982). In non-pathogenic bacteria, *e.g. Streptomyces*, ChOx is secreted and is a component of the metabolic pathway for utilizing cholesterol as a carbon source (Fukuda *et al*, 1973; Cheetman *et al*, 1982). Pathogenic bacteria, *e.g. Rhodococcus equi* and slow-growing *Mycobacterium* spp., require ChOx for invasion of the host macrophage and cholesterol regulates the expression of this enzyme (Fernanandez-Garayzabal *et al*, 1996; Navas *et al*, 2001). ChOx is a water soluble, interfacial enzyme that binds transiently to the membrane surface during catalysis (Sampson *et al*, 1998).

ChOxs utilise a Ping-Pong mechanism with respect to the steroid substrate and O_2. In the reductive half-reaction the cholesterol reduces FAD by hydride transfer from the 3C of cholesterol to N5 of the flavin moiety yielding cholest-5-ene-3-one

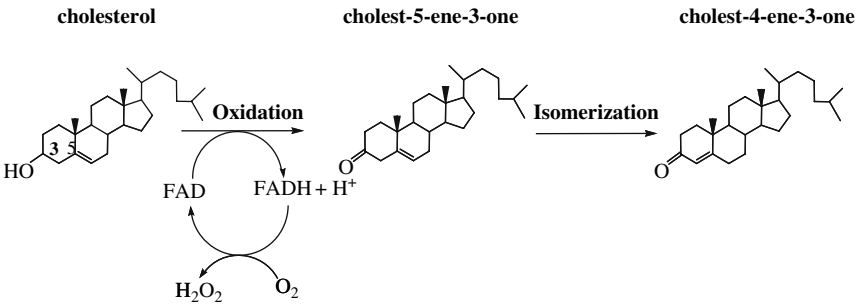


Figure 4. Cholesterol oxidation and isomerization catalysed by cholesterol oxidase

and 1,5-dihydroflavin ($FADH_2$) (Medina *et al*, 1997). In the oxidative half-reaction the reduced flavin ($FADH_2$) is converted back to the active, oxidized form (FAD) by O_2 with the generation of H_2O_2.

2.4. Catalases

Catalases (hydrogen-peroxide:hydrogen-peroxide oxidoreductase, CAT, EC 1.11.1.6.), more correctly hydroperoxidases, catalyse the degradation of H_2O_2 to H_2O and O_2 : $2H_2O_2 \rightarrow 2H_2O + O_2$.

Considering the possibility of industrial enzyme production, a variety of different micro-organisms including bacteria (*e.g. Micrococcus, Bacillus, Microscilla, Alcaligenes* spp.), moulds (*e.g. Aspergillus, Penicillium, Thermomyces, Thermoascus, Acremonium* spp.) and yeasts (*e.g. Saccharomyces, Candida, Mycotorula* spp.) are known as good CAT producers. CATs from animal sources (*e.g.* bovine liver) are generally cheap, therefore the production of microbial CATs will only be economical when better producer (preferably recombinant) strains and cheap technology can be used or CATs with special properties (*e.g* working at high or low temperatures or at alkaline or acidic pH) are produced (Fusho and Yajima, 1995; Kou *et al*, 1998; Takeuchi and Isobe, 1999).

In general, three distinct types of proteins, which do not share either sequence or structural homologies, exhibit significant CAT activity. The class I CATs, which are the most widely spread in nature, consists of monofunctional haem-containing enzymes which can be divided into two subclasses having large (> 75KDa) or small (< 60 kDa) subunits (Klotz *et al*, 1997). The second, less widespread class includes bifunctional haem-containing CAT-peroxidases, which are closely related to plant peroxidases. The third class comprises members of the non-haem or Mn-containing CATs (Wu *et al*, 2004).

Monofunctional haem-containing CATs all possess a two-stage mechanism for the degradation of H_2O_2. The first, oxidative, step in the catalyses is the mono-oxygen transfer from H_2O_2 to the iron centre with the release of water (Fig. 5). Under physiological conditions the resting state of haem CATs contains iron (III). In the oxidation, one molecule of H_2O_2 oxidizes the haem to an oxy-ferryl species, in which one electron is removed from the iron and one electron is removed from the coordinated porphyrin ring to generate a porphyrin radical cation (Fig. 5). In the second stage, another H_2O_2 molecule is utilised to regenerate the enzyme by reducing the porphyrin radical coordinated oxy-ferryl species to the resting haem

$$Enz(Por–Fe^{III}) + H_2O_2 \longrightarrow Enz(Por^{+\bullet}–Fe^{IV}=O) + H_2O \quad (1)$$

$$Enz(Por^{+\bullet}–Fe^{IV}=O) + H_2O_2 \longrightarrow Enz(Por–Fe^{III}) + H_2O + O_2 \quad (2)$$

Figure 5. The two-stage process of H_2O_2 disproportionation catalysed by catalases

releasing H_2O and O_2 (Matsunaga and Shird, 2004). The haem reactivity is enhanced by the axial phenolate ligand of Tyr.

CATs do not follow classical Michaelis-Menten kinetics except across a certain range of substrate concentrations (Switala and Loewen, 2002). Most small subunit enzymes are inactivated by H_2O_2 at concentrations above 300–500 mM and never reach the maximum rate extrapolated from the Michaelis–Menten curve determined at lower substrate concentrations. Large subunit enzymes show a greater resistance to H_2O_2 damage even at concentrations above 3 M and the observed rates exceed the theoretical Michaelis-Menten maximum rate (Switala and Loewen, 2002).

Most of the monofunctional haem-containing CATs are uniquely stable and resist proteolysis due to their very rigid structure that resists unfolding (Fig. 6). Enzymes with larger subunit such as HPII CAT from *E. coli*, show enhanced thermal stability, the activity of the enzyme starting to drop at temperatures above 80 °C (Switala *et al*., 1999).

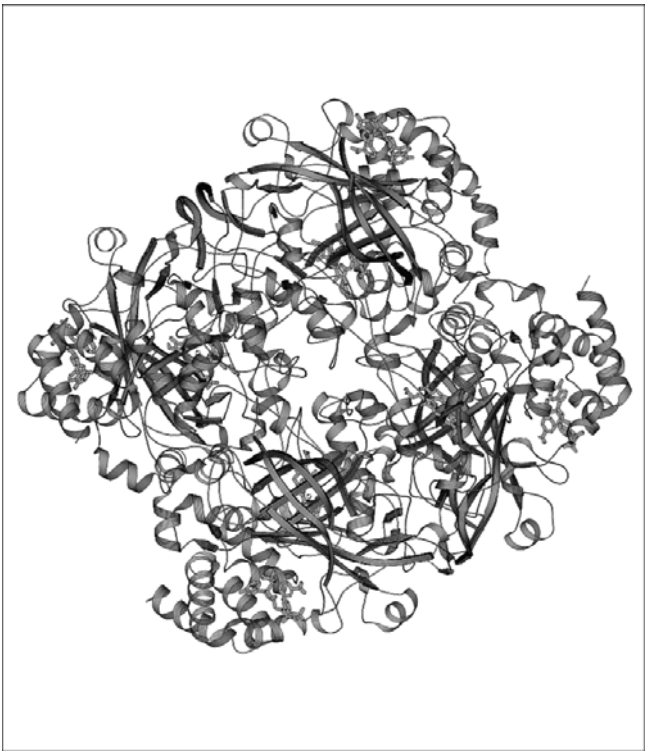


Figure 6. The tetrameric structure of catalase from beef liver. Each domain contains haem and NADP cofactor represented with ribbon diagram (Kraulis, 1991)

3. BIOSENSORS BASED ON OXIDASES AND CATALASE

3.1. Biosensors Based on Glucose Oxidase

Currently, GOx is the most important model enzyme in both basic and applied biosensor research because its enzymological properties are well understood and it is very cheap in comparison to other oxidases. Different design approaches have been considered and employed in the construction of GOx-based biosensors, including needle type sensors, sensors for *in vivo* monitoring, flow injection analysis (FIA) applications, combinations of glucose biosensors with microdialysis sampling techniques and disposable sensors (Wilson and Gifford, 2005).

Electrochemical glucose biosensors prepared by immobilising GOx on the surface of carbon film electrodes using glutaraldehyde and bovine serum albumin (BSA) with or without Nafion (a sulfonated tetrafluorethylene copolymer) have been reported by Florescu and Brett (2005). An ISFET (Ion Sensitive Field Effect Transistor) biosensor containing GOx and horseradish peroxidase (HRP) co-immobilized with BSA using glutaraldehyde and covered by poly(4-vinylpyridine-co-styrene) polymeric film could also be used for glucose determination (Volotovsky and Kim, 1998). Monolayer enzyme electrodes have been made by covalently or electrostatically binding the recognition molecule onto electrodes modified with conducting polymer films (*e.g.* self-assembled monolayers, Langmuir–Blodgett films; Malhotra and Singhal, 2003). A flow-through electrochemical micro-cell can be used as an on-line detector in microdialysis-based assays. The micro-detector modified by enzyme-based chemistry can be directly connected to the outlet of the microdialysis probe (Gáspár *et al*., 2004).

Redox couples, or mediators, are able to shuttle electrons between the redox centre of the enzyme and the electrode. The electron transfer mediator must be chemically stable in both the reduced and oxidised forms, have a relatively low and stable redox potential, and should be easy to immobilize at the electrode surface (Harwood and Pouton, 1996). Carbon paste electrodes or graphite–Teflon rigid composite biosensors offer the possibility of co-immobilization of several enzymes *via* simple physical inclusion in the bulk of the electrode matrix by non-covalent linkages in the presence of a mediator like ferrocene (Guzman-Vázquez de Prada *et al*., 2004). Yu *et al*. (2003) constructed a biosensor by immobilizing GOx with titanium isopropoxide forming GOx-titania sol-gel film, which retains the native structure and activity of the entrapped enzymes. Carbon film resistor electrodes can be modified with Prussian Blue (PB, ferric hexacyanoferrate) and then covered with a layer of covalently immobilized enzyme. These enzyme electrodes can be used to detect glucose *via* the oxidation of H_2O_2 at +50mV *vs*. Ag/AgCl in the low micromolar range, while also avoiding or reducing electrochemical interference (Ricci and Palleschi, 2005). Electrodes produced with the 'screen printing' thick-film technique can be chemically modified with PB prior to enzyme immobilisation (Newman and Turner, 2005).

The use of metallized carbon electrodes resulted in remarkably selective amperometric biosensors because these transducers eliminated major electroactive

interference. Platinization has been used to increase electrode surface area and thus electrode sensitivity, immobilizing GOx by the electropolymerization of *o*-phenylenediamine (Reyes De Corcuera *et al*., 2005). Glucose nanosensors based on the co-electrodeposition of an osmium redox polymer/enzyme composite on low-noise carbon fibre nanoelectrodes have been constructed (Fei *et al*., 2005). Chromium and manganese half-sandwich complexes were developed as mediators for GOx since their sizes are similar to that of ferrocene derivatives and they contain a π-ligand for interaction with the enzyme co-factor (Forrow and Walters, 2004). Screen-printed electrodes modified with ruthenium dioxide were investigated as amperometric sensors to immobilize GOx onto the electrode surface using Nafion films (Kotzian *et al*., 2005).

Recently there has been considerable interest in the development of single-use sensor strips especially in the fields of decentralized medical diagnostics or on-site environmental monitoring (Wang *et al*., 1995). These strips can be regarded as disposable electrochemical cells onto which a sample droplet is placed. A miniaturized total analytical system (μTAS) for on-line monitoring of glucose in biological systems based on integrated microdialysis sampling and photolithographically-prepared glucose electrodes has been developed (Wang, 1999).

Interest in optical fibre biosensors is growing fast since they confer several advantages compared to biochemical and clinical analyses due to their simple and flexible construction. Malhotra *et al*. (2005) presented a glucose biosensor in which GOx was immobilized onto poly-3-hexyl thiophene mixed with stearic acid (P3HT/SA) LB films. The activity of GOx-containing LB films was measured by a colorimetric method using *o*-dianisidine as a chromogenic dye with absorbance at 540 nm. Capillary electrophoresis integrated with an optical biosensor was built by modifying the detector part of the electrophoresis equipment with the redox-sensitive polymer film polyaniline (PANI) containing immobilized GOx (Bossi *et al*., 2003). Measuring molecular fluorescence is a very promising technique in biosensing because of its high sensitivity. Fibre-optic glucose sensors based on this phenomenon have been developed by immobilizing GOx onto an oxygen optrode composed of either decacyclene in silicone (ex 385 nm, em 450–600 nm) or the ruthenium complex tris(1,10-phenanthrolene)ruthenium chloride (ex 447 nm, em 604 nm). The formation of hydrogen peroxide can be followed by measuring chemiluminescence (luminol) or *via* the formation of a fluorescent oxidation product from non-fluorescent *p*-hydroxyphenyl acetic acid (HPA) or homovanillic acid (HVA) in the presence of HRP. A decrease in the pH in the enzyme layer can be determined taking advantage of the light absorption (*e.g.* bromocresol green) or fluorescence (*e.g.* hydroxypyrenetrisulphonate) of acid/base indicators. In a different type of glucose biosensor GOx itself (the prosthetic group, FAD), or GOx chemically modified with a fluorescein derivative (GOx-FS), served as an indicator for recording changes in fluorescence (Pickup *et al*., 2005).

3.2. Multienzyme Biosensors Based on Glucose Oxidase

Multienzyme systems involving GOx have been developed for the rapid, reproducible and cheap determination of different disaccharides and polysaccharides. For example, the concentration of maltose was determined using a bienzyme cell in which amyloglucosidase and GOx were co-immobilized on a protein membrane using glutaraldehyde followed by amperometric detection. For lactose sensing, β-galactosidase, GAO and GOx enzymes were co-immobilized either on the surface of a measuring electrode using a triacetate cellulose membrane, or on the surface of different resins, or on the controlled-pore glass used in analytical reactors, or on protein membranes by covalent coupling (Adányi *et al*., 1999). Sensors for total D-glucose ($\alpha+\beta$) and sucrose have been developed using pectin as a novel matrix to enhance enzyme entrapment and stabilisation on rhodinised carbon electrodes. Total D-glucose can be measured by enzymatically transforming all α-glucose anomers to the β form using mutarotase. For sucrose detection, a multienzyme system involving invertase, mutarotase and GOx were used (Jawaheer *et al*., 2003). With the incorporation of lysozyme during the immobilization step, considerable enhancement of the operational stability of a biosensor has been demonstrated in the case of sucrose determination (Gouda *et al*., 2002). Marconi *et al*. (2004) developed a method for the determination of gelatinised starch in processed cereal foods by co-immobilization of amyloglucosidase and GOx on a Pt electrode surface while the third enzyme, α-amylase, was added to the solution under analysis. Wu *et al*. (2005) proposed a biosensor for the determination of glucosinolates employing a bienzyme system involving myrosinase and GOx co-immobilized onto an eggshell membrane on the surface of a dissolved oxygen electrode.

3.3. Biosensors Based on Galactose Oxidase

GAO is used in biosensors developed for the determination of galactose or dihydroxyacetone in different biofermentation processes, or for measuring lactose primarily in food samples. Schumacher *et al*. (1994) constructed an electrode for the determination of galactose and galactose-containing disaccharides in which GAO was immobilized in gelatine between two dialysis membranes and stuck to an electrode. Malhotra *et al*. (2005) prepared an amperometric biosensor to measure galactose in milk and blood serum by immobilizing the enzymes in Langmuir–Blodgett (LB) films of poly-3-hexyl thiophene (P3HT) mixed with stearic acid (SA), and depositing the LB film onto indium tin-oxide (ITO) coated glass plates. Hasebe and Uchiyama (2000) demonstrated that the substrate specificity of GAO from *F. dendroides* changed dramatically in the presence of L-histidine resulting in catalysis of the air oxidation of L-ascorbate. Using this observation a flow-type system was developed for the indirect measurement of histidine. Since the catalytic activity of GAO is inhibited by free radicals a biosensor based on enzyme inhibition was developed to detect superoxide and nitrogen oxide radicals using GAO entrapped in a gel-like κ-carrageenan membrane coupled to an amperometric oxygen electrode (Campanella *et al*., 2000).

3.4. Biosensors Based on Cholesterol Oxidase

To measure free and total cholesterol contents in biological samples, ChOx-based biosensors and bienzyme cells containing immobilized cholesterol esterase (ChEt) and ChOx were developed, respectively. Ram *et al.* (2001) reported a biosensor in which ChOx and ChEt were bound to either a collagen membrane or to conducting polymer matrices. The electrochemical redox processes taking place in enzyme-layered films deposited either on platinum or ITO-coated glass plates have been investigated. Gobi and Mizutani (2001) constructed a direct amperometric biosensor by layer-by-layer nanothin film formation using ChOx and poly(styrenesulfonate) on a monolayer of HRP covalently immobilized on Au-alkanethiolate electrodes. ChOx and ChEt were entrapped within polypyrrole (PPy) films on a platinum disc electrode during electrochemical polymerisation (Singh *et al.*, 2004). Situmorang *et al.* (1999) studied the conversion of cholesterol esters by flow injection potentiometry using a tungsten electrode and monitoring ferricyanide/ferrocyanide conversion. Hall and Turner (1991) described an amperometric organic-phase enzyme electrode (OPEE) using ChOx to determine the concentration of cholesterol in a chloroform/hexane mixture. Pena *et al.* (2001) reported on a bienzyme amperometric composite biosensor for the determination of free and total cholesterol in food samples. ChOx and HRP together with potassium ferrocyanide as a mediator were incorporated into a graphite-Teflon (30%–70%) matrix. The enzyme membrane was prepared from ChEt and ChOx using a photosensitive polymer and an ultra-thin dialysis membrane, and this was applied in a flow system with a Clark-type oxygen electrode (Endo *et al.*, 2003). An amperometric cholesterol biosensor using carbon nanotubes has been produced by layer-by-layer deposition of a cationic polyelectrolyte (poly(diallyldimethylammonium) chloride; PDDA) and ChOx on a multi-walled carbon nanotube-modified gold electrode (Guo *et al.*, 2004). The hybrid composite film was used to immobilize ChOx on the surface of PB-modified glass carbon electrode (Tan *et al.*, 2005).

The FIA method was investigated by covalent co-immobilization of enzymes to the silica in packed-bed reactors (IMMER) and using an amperometric HRP electrode or photometric determination of H_2O_2 (Baticz and Tömösközi, 2002). Optical cholesterol biosensors have also been constructed by immobilizing ChOx and octadecyl silica (ODS) particles in either hydrogel network matrices of copolymers of poly(vinyl alcohol)/hydroxyethyl carboxymethyl cellulose (PVA/HECMC) or sol-gel (Wu and Choi, 2003).

3.5. Biosensors Based on Catalase

CAT decomposes H_2O_2 and remains active when kept in appropriate organic solvent environments. Wang *et al.* (1995) demonstrated the applicability of CAT for organic-phase biosensing by immobilising the enzyme onto a glassy carbon amperometric transducer by casting a mixed enzyme/Eastman-AQ polymer solution. Horozova *et al.* (2002) prepared an OPEE by immobilizing the enzyme within a

polymeric film on a spectrographic graphite electrode. Salimi *et al.* (2005) investigated the direct voltammetry and electrocatalytic properties of CAT adsorbed onto the surface of multi-wall carbon nanotubes. Collagen film was formed on pyrolytic graphite electrodes, which provided a suitable microenvironment for haemoglobin and CAT to transfer electrons directly to the underlying electrodes (Li *et al.*, 2005). Campanella *et al.* (2001) developed an OPEE by immobilizing CAT in a κ-carrageenan gel and demonstrated its use in chloroform, toluene, chlorobenzene and ethyl acetate in a stirred reactor for monitoring the hydrogen peroxide content of extra virgin olive oil during an artificial rancidification process. Joo *et al.* (1996) enhanced enzyme stability by mixing CAT with polyethylene glycol (PEG), however, the response was dependent upon the molecular weight and concentration of PEG. A quick analytical method was developed to indirectly monitor the water (activator) content of various butter and margarine samples using CAT immobilized by glutaraldehyde on a natural protein membrane in a thin-layer enzyme cell and connected to a stopped-flow injection analyser (SFIA) system with an amperometric detector (Adányi and Váradi, 2004).

4. OTHER APPLICATIONS

4.1. Food Microbiology and Sanitizing

The well-studied bactericide effect of the GOx/glucose system by virtue of H_2O_2 production may be exploitable in the food industry to combat food-poisoning microbes (Dobbenie *et al.*, 1995; Fuglsang *et al.*, 1995; Massa *et al.*, 2001). It is noteworthy that GOx affects the growth of fungi (*e.g. Candida* and *Aspergillus* spp.) to very different extents clearly depending on the inherent antioxidant potentials of the target organisms (Leiter *et al.*, 2004). Combinations of the H_2O_2-generating and H_2O_2-consuming enzymes GOx and CAT are also deleterious for aerobic microorganisms because of the fast depletion of available oxygen and hence can be used as antimicrobial agents (Dondero *et al.*, 1993; Fuglsang *et al.*, 1995; Leiter *et al.*, 2004). The O_2-scavenging effect of the GOx/CAT system has potential food-related applications in the control of lipid peroxidation (*e.g.* in mayonnaises; Isaksen and Adler-Nissen, 1997), enzymatic and non-enzymatic browning (*e.g.* in wine and apple and pear purées; Ough, 1975; Parpinello *et al.*, 2002), off-odour and off-flavour development (*e.g.* in mayonnaises; Jafar *et al.*, 1994) and colour changes (*e.g.* in grape juice; Castellari *et al.*, 2000).

The antimicrobial potential of GOx is exploited most effectively in combination with peroxidases such as lactoperoxidases (Sandholm *et al.*, 1988; Popper and Knorr, 1997; Revol-Junelles *et al.*, 2001), HRP (Hill *et al.*, 1997) and chloroperoxidases (Jones *et al.*, 1998). GOx alone or encapsulated into liposomes together with peroxidases may be used for the removal and disinfection of bacterial and fungal biofilms from medical implants and in the oral cavity (Johansen *et al.*, 1997; Jones *et al.*, 1998). Multienzymatic system encapsulation techniques are developing fast. For example, microbeads containing all the components of the GOx-lactoperoxidase

system (enzymes, glucose, thiocyanate) are now available, which may facilitate the use of this system as a preservative for raw milk especially in developing countries (Iacquot and Poncelet, 2003). CAT in immobilized form can also be used to remove excess H_2O_2 after 'cold pasteurisation' (*i.e.* milk treated with H_2O_2 at a concentration of 0.05–0.25%), which might be useful and applicable in areas of tropical or hot climates (Tarhan, 1995; Akertek and Tarhan, 1995).

The increasing interest in and consumer demand for 'low-alcohol' wines has triggered research into the optimisation of glucose conversion to gluconic acid in grape must by GOx (Pickering *et al*., 1998, 1999) and the introduction of recombinant *Saccharomyces cerevisiae* strains into the winemaking process that harbour the *A. niger* GOx gene (Malherbe *et al*., 2003). The use of GOx in winemaking also improves wine stability (Gomez *et al*., 1995) and alcohol acid balance (Pickering *et al*., 1998). An additional application for GOx is in the stabilization of beer (Ohlmeyer, 1957).

GOx has been employed in baking to improve the strength of the gluten network *via* H_2O_2-triggered permanent cross-linking of wheat dough proteins (Hilhorst *et al*., 1999; Dunnewind *et al*., 2002; Primo-Martin *et al*., 2004; Rasiah *et al*., 2005), and to reduce the stickiness of enzyme-supplemented doughs (Matinez-Anaya and Jimenez, 1998; Collar *et al*., 1998). The performance of GOx is comparable to that of chemical oxidants such as the bromates traditionally used in bakeries (Vemulapalli *et al*., 1998). Combination of GOx with xylanase and α-amylase further improves dough and bread quality, *e.g.* dough stability and extensibility and bread volume (Indrani *et al*., 2003; Primo-Martin *et al*., 2003). GOx may also catalyse the formation of arabinoxylane-arabinoxylane cross-links and xylanase may compensate this disadvantageous effect by cleaving large arabinoxylan complexes (Primo-Martin *et al*., 2005).

4.2. Textile Industry

The industrial application of CAT is wide-ranging because its substrate, H_2O_2, is frequently used as a strong biocide, oxidant and bleaching agent in the textile, pulp and paper, wood, food, *etc*. industries (Galante and Formantici, 2003). Bleaching with H_2O_2 in the textile industry is performed after desizing and scouring, but before dying. CAT is now used to decompose excess H_2O_2 and thus contributes to reducing water consumption (no repeated rinsing), the avoidance of loading wastewaters with pollutants (like derivatives of $Na_2S_2O_3$ or $KMnO_4$), and it enables the dying process to be performed directly in CAT-treated rinsing baths (Amorim *et al*., 2002; Galante and Formantici, 2003). To avoid unacceptable interference by proteins (CATs) in the dyeing process (Tzanov *et al*., 2001), the introduction of immobilized CATs (Costa *et al*., 2002; Opwis *et al*., 2004) or, alternatively, immobilized thermo-alkali-stable bacterial CAT-peroxidases (Paar *et al*., 2001; Fruhwirth *et al*., 2002) are recommended for the treatment and recycling of textile bleaching effluents for dyeing. Two other interesting possibilities for technological innovation in this field are the application of immobilized GOx to generate H_2O_2 for

textile bleaching (Tzanov *et al*., 2002) and immobilized CAT-peroxidase producer alkalothermophilic *Bacillus* cells to decompose excess H_2O_2 in bleaching effluents (Paar *et al*., 2003).

REFERENCES

Adányi, N., Szabó, E.E. and Váradi, M. (1999) Multi-enzyme biosensors with amperometric detection of lactose in milk and dairy products. Z. Lebensm. Unters.-Forsch. A. *209*, 220–226.

Adányi, N. and Váradi, M. (2004) Catalase-based thin-layer enzyme cell used in organic-phase FIA system for determination of moisture in oily foods. Eur. Food Res. Technol. *219*, 432–437.

Akertek, E. and Tarhan, L. (1995) Characterization of immobilized catalases and their application in pasteurization of milk with H_2O_2. Appl. Biochem. Biotech. *50*, 291–303.

Alvarez-Icaza, M., Kalisz, H.M., Hecht, H.J., Aumann, K.D., Schomburg, D. and Schmid, R.D. (1995) The design of enzyme sensors based on the enzyme structure. Biosens. Bioelectron. *10*, 735–742.

Amorim, A.M., Gasques, M.D.G., Andreaus, J. and Scharf, M. (2002) The application of catalase for the elimination of hydrogen peroxide residues after bleaching of cotton fabrics. Anais Acad. Brasil. Ciències *74*, 433–435.

Baticz, O. and Tömösközi, S. (2002) Determination of total cholesterol content in food by flow injection analysis with immobilized cholesterol oxidase enzyme reactor. Nahrung *46*, 46–50.

Bossi, A., Castelletti, L., Piletsky, S.A. and Righetti, P.G. (2003) Towards the development of an integrated capillary electrophoresis optical biosensor. Electrophoresis *24*, 3356–3363.

Bourdillon, C., Demaille, C., Moiroux, J. and Saveant, J.M. (1993) New insights into the enzymic catalysis of the oxidation of glucose by native and recombinant glucose oxidase mediated by electrochemically generated one-electron redox cosubstrates. J. Am. Chem. Soc. *115*, 1–10.

Bretting, H. and Jacobs, G. (1987) The reactivity of galactose oxidase with snail galactans, galactosides and D-galactose-composed oligosaccharides. Biochim. Biophys. Acta. *913*, 342–348.

Campanella, L., Persi, L. and Tomassetti M. (2000) A new tool for superoxide and nitric oxide radicals determination using suitable enzymatic sensors. Sensor. Actuat. B-Chem. *68*, 351–359.

Campanella, L., Sammartino, M.P., Tomassetti, M. and Zannella S., (2001) Hydroperoxide determination by a catalase OPEE: application to the study of extra virgin olive oil rancidification process. Sensor. Actuat. B-Chem. *76*, 158–165.

Castellari, M., Matricardi, L., Arfelli, G., Carpi, G. and Galassi, S. (2000) Effects of high hydrostatic pressure processing and of glucose oxidase-catalase addition on the color stability and sensorial score of grape juice. Food Sci.Technol. Int. *6*, 17–23.

Chan, T.W. and Bruice, T.C. (1977) One and two electron transfer reactions of glucose oxidase. J. Am. Chem. Soc. *99*, 2387–2389.

Cheetham, P.S., Dunnill, P. and Lilly, M.D. (1982) The characterization and interconversion of three forms of cholesterol oxidase extracted from *Nocardia rhodochrous*. Biochem. J. *201*, 515–521.

Chelikani, P., Fita, I. and Loewen, P.C. (2004) Diversity of structures and properties among catalases. Cell. Mol. Life Sci. *61*, 192–208.

Collar, C., Andreu, P. and Martinez-Anaya, M.A. (1998) Interactive effects of flour, starter and enzyme on bread dough machinability. Z. Lebensm. Unters.-Forsch. A. *207*, 133–139.

Costa, S.A., Tzanov, T., Carneiro, F., Gübitz, G.M. and Cavaco-Paulo, A. (2002) Recycling of textile bleaching effluents for dyeing using immobilized catalase. Biotechnol. Lett. *24*, 173–176.

Crueger, A. and Crueger, W. (1990) Glucose transforming enzyme. In: Microbial Enzymes and Biotechnology (Fogarty, W.M. and Kelly, C.T. Eds.) Elsevier Applied Science, London, pp. 177–226.

Dobbenie, D., Uyttendaele, M. and Debevere, J. (1995) Antibacterial activity of the glucose oxidase/glucose system in liquid whole egg. J. Food Protect. *58*, 273–279.

Dondero, M., Egaña, W., Tarky, W., Cifuentes, A. and Torres, J.A. (1993) Glucose oxidase/catalase improves preservation of shrimp (*Heterocarpus reedi*). J. Food Sci. *58*, 774–779.

Dunnewind, B., van Vliet, T. and Orsel, R. (2002) Effect of oxidative enzymes on bulk rheological properties of wheat flour doughs. J. Cereal Sci. *36*, 357–366.

Endo, H., Maita, M., Takikawa, M., Ren, H., Hayashi, T., Urano, N. and Mitsubayashi, K. (2003) Enzyme sensor system for determination of total cholesterol in fish plasma. Fisheries Sci. *69*, 1194–1199.

Fei, J.J., Wu, K.B., Wang, F. and Hu, S.S. (2005) Glucose nanosensors based on redox polymer/glucose oxidase modified carbon fiber nanoelectrodes. Talanta *65*, 918–924.

Fernanandez-Garayzabal, J.F., Delgado, C., Blanco, M.M., Suarez, G. and Dominguez, L. (1996) Cholesterol oxidase from *Rhodococcus equi* is likely the major factor involved in the cooperative lytic process (CAMP reaction) with *Listeria monocytogenes*. Lett. Appl. Microbiol. *22*, 249–252.

Florescu, M. and Brett C.M.A. (2005) Development and evaluation of electrochemical glucose enzyme biosensors based on carbon film electrodes. Talanta *65*, 306–312.

Forrow, N.J., Sanghera, G.S. and Walters, S.J. (2002) The influence of structure in the reaction of electrochemically generated ferrocenium derivatives with reduced glucose oxidase. Dalton Trans. *16*, 3187–3194.

Forrow, N.J. and Walters, S.J. (2004) Transition metal half-sandwich complexes as redox mediators to glucose oxidase. Biosens. Bioelectron. *19*, 763–770.

Fruhwirth, G.O., Paar, A., Gudelj, M., Cavaco-Paulo, A., Robra, K.H. and Gübitz, G.M. (2002) An immobilized catalase peroxidase from the alkalothermophilic *Bacillus* SF for the treatment of textile-bleaching effluents. Appl. Microbiol. Biotechnol. *60*, 313–319.

Fuglsang, C.C., Johansen, C., Christgau, S. and Adler-Nissen, J. (1995) Antimicrobial enzymes: applications and future potential in the food industry. Trends Food Sci. Tech. *6*, 390–396.

Fukuda, H., Kawakami, Y. and Nakamura, S. (1973) A method to screen anticholesterol substances produced by microbes and a new cholesterol oxidase produced by *Streptomyces violascens*. Chem. Pharm. Bull. (Tokyo) *21*, 2057–2060.

Fusho, Y. and Yajima, Y. (1995) Catalase and production process thereof. Patent No.: EP0663405.

Fusho, Y. and Yajima, Y. (1997) Catalase from *Bacillus* and process for producing the same. Patent No.: US5622849.

Galante, Y.M. and Formantici, C. (2003) Enzyme applications in detergency and in manufacturing industries. Curr. Org. Chem. *7*, 1399–1422.

Gáspár, S, Wang, X., Suzuki, H. and Csöregi, E. (2004) Amperometric biosensor-based flow-through microdetector for microdialysis applications. Anal. Chim. Acta *525*, 75–82.

Giordano, R.S., Bereman, R.D., Kosman, D.J. and Ettinger, M.J. (1974) Stereoelectronic properties of metalloenzymes. II. Effects of ligand coordination on the electron spin resonance spectrum of galactose oxidase as a probe of structure and function. J. Am. Chem. Soc. *96*, 1023–1026.

Gobi, K.V. and Mizutani, F. (2001). Layer-by-layer construction of an active multilayer enzyme electrode applicable for direct amperometric determination of cholesterol. Sensor. Actuat. B-Chem. *80*, 272–277.

Gomez, E., Martinez, A. and Laencina, J. (1995) Prevention of oxidative browning during wine storage. Food Res. Int. *28*, 213–217.

Goodhue, C.T. and Risley, H.A. (1978) Production of cholesterol oxidase by *Nocardia cholesterolicum.* Patent No: US3909359.

Gouda, M.D., Kumar, M.A., Thakur, M.S. and Karanth, N.G. (2002). Enhancement of operational stability of an enzyme biosensor for glucose and sucrose using protein based stabilizing agents. Biosens. Bioelectron. *17*, 503–507.

Gouda, M.D., Singh, S.A., Rao, A.G., Thakur, M.S. and Karanth, N.G. (2003) Thermal inactivation of glucose oxidase. Mechanism and stabilization using additives. J. Biol. Chem. *278*, 24324–24333.

Goudsmit, E.M., Matsuura, F. and Blake, D.A. (1984) Substrate specificity of D-galactose oxidase. Evidence for the oxidation of internally linked galactosyl residues of *Helix pomatia* galactogen. J. Biol. Chem. *259*, 2875–2878.

Guo, M.L., Chen, J.H., Li, J., Nie, L.H. and Yao, S.Z. (2004) Carbon nanotubes-based amperometric cholesterol biosensor fabricated through layer-by-layer technique. Electroanalysis *16*, 1992–1998.

Guzman-Vázquez de Prada, A., Pena, N., Parrado, C., Reviejo, A.J. and Pingarrón, J.M. (2004) Amperometric multidetection with composite enzyme electrodes. Talanta *62*, 896–903.

Hall, G.F. and Turner, A.P.F. (1991) An organic phase enzyme electrode for cholesterol. Anal. Lett. *24*, 1375–1388.

Harwood, G.W.J. and Pouton, C.W. (1996) Amperometric enzyme biosensors for metabolites the analysis of drugs and metabolites. Adv. Drug Deliver. Rev. *18*, 163–191.

Hasebe, Y. and Uchiyama, S. (2000). Amperometric flow-type L-histidine sensor using an immobilized galactose oxidase reactor, based on a novel catalytic activity induced by exogenous histidine. Sensor. Actuat. B-Chem. *66*, 12–15.

Hilhorst, R., Dunnewind, B., Orsel, R., Stegeman, P., van Vliet, T., Gruppen, H. and Schols, H.A. (1999) Baking performance, rheology, and chemical composition of wheat dough and gluten affected by xylanase and oxidative enzymes. J. Food Sci. *64*, 808–813.

Hill, K.J., Kaszuba, M., Creeth, J.E. and Jones, M.N. (1997) Reactive liposomes encapsulating a glucose oxidase-peroxidase system with antibacterial activity. Biochim. Biophys. Acta *1326*, 37–46.

Horozova, E., Dimcheva, N. and Jordanova Z. (2002) Study of catalase electrode for organic peroxidase assays. Bioelectrochemistry *58*, 181–187.

Indrani, D., Prabhasankar, P., Rajiv, J. and Rao, G.V. (2003) Scanning electron microscopy, rheological characteristics, and bread-baking performance of wheat-flour dough as affected by enzymes. J. Food Sci. *68*, 2804–2809.

Inouye, Y., Taguchi, K., Fuji, A., Ishimaru, K., Nakamura, S. and Nomi, R. (1982) Chem. Pharm. Bull. (Tokyo) *30*, 951–958.

Isaksen, A. and Adler-Nissen, J. (1997) Antioxidative effect of glucose oxidase and catalase in mayonnaises of different oxidative susceptibility. I. Product trials. LWT-Food Sci. Technol. *30*, 841–846.

Ito, N., Phillips, S.E., Stevens, C., Ogel, Z.B., McPherson, M.J., Keen, J.N., Yadav, K.D. and Knowles, P.F. (1991) Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase. Nature *350*, 87–90.

Jacquot, M. and Poncelet, D. (2003) Multienzymic system encapsulation: application to the lactoperoxidase system. J. Chem. Chem. Eng. Tech. *12*, 581–584.

Jafar, S.S., Hulton, H.O., Bimbo, A.P., Crowther, J.B. and Barlow, S.M. (1994) Stabilization by antioxidants of mayonnaise made from fish oil. J. Food Lipids *1*, 295–311.

Jawaheer, S., White, S.F, Rughooputh, S.D.D.V. and Cullen D.C. (2003) Development of a common biosensor format for an enzyme based biosensor array to monitor fruit quality. Biosens. Bioelectron. *18*, 1429–1437.

Johansen, C., Falholt, P. and Gram, L. (1997) Enzymatic removal and disinfection of bacterial biofilms. Appl. Environ. Microbiol. *63*, 3724–3728.

Jones, M.N., Hill, K.J., Kaszuba, M. and Creeth, J.E. (1998) Antibacterial reactive liposomes encapsulating coupled enzyme systems. Int. J. Pharm. *162*, 107–117.

Joo, H., Yoo, Y.J. and Ryu, D.D.Y. (1996) A biosensor stabilized by polyethylene glycol for monitoring of hydrogen peroxide in organic solvent media. Enzyme Microb. Tech. *19*, 50–56.

Klibanov, A.M., Alberti, B.N. and Marletta, M.A. (1982) Stereospecific oxidation of aliphatic alcohols catalyzed by galactose oxidase. Biochem. Biophys. Res. Commun. *108*, 804–808.

Klotz, M.G. and Klassen, G.R. and Loewen, P.C. (1997) Phylogenetic relationships among prokaryotic and eukaryotic catalases. Mol. Biol. Evol. *14*, 951–958.

Kona, R.P., Qureshi, N. and Pai, J.S. (2001) Production of glucose oxidase using *Aspergillus niger* and corn steep liquor. Bioresource Technol. *78*, 123–126.

Kotzian, P., Brázdilová, P., Kalcher, K. and Vytras, K. (2005). Determination of hydrogen peroxide, glucose and hypoxanthine using (bio)sensors based on ruthenium-dioxide modified screen-printed electrodes. Anal. Lett. *38*, 1099–1113.

Kou, S., Tagoyama, Y., Yamazumi, T. and Yamamoto, H. (1998) Catalase and its production. Patent No.: JP10257883.

Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Cryst. *24*, 946–950.

Leiter, É., Marx, F., Pusztahelyi, T., Haas, H. and Pócsi, I. (2004) *Penicillium chrysogenum* glucose oxidase - a study on its antifungal effects. J. Appl. Microbiol. *97*, 1201–1209.

Malherbe, D.F., du Toit, M., Cordero Otero, R.R., van Rensburg, P. and Pretorius, I.S. (2003) Expression of the *Aspergillus niger* glucose oxidase gene in *Saccharomyces cerevisiae* and its potential applications in wine production. Appl. Microbiol. Biot. *61*, 502–511.

Malhotra, B.D. and Singhal, R. (2003) Conducting polymer based biomolecular electronic devices Pramana. J. Phys. *61*, 331–343.

Malhotra, B.D., Singhal, R., Chaubey, A., Sharma, S.K. and Kumar, A. (2005) Recent trends in biosensors. Curr. Appl. Phys. *5*, 92–97.

Marconi, E., Messia, M.C., Palleschi, G. and Cubadda, R. (2004) Maltose biosensor for determining gelatinized starch in processed cereal foods. Cereal Chem. *81*, 6–9.

Martinez-Anaya, M.A. and Jimenez, T. (1998) Physical properties of enzyme-supplemented doughs and relationship with bread quality parameters. Z. Lebensm. Unters.-Forsch. A. *206*, 134–142.

Massa, S., Petruccioli, M., Brocchi, G.F., Altieri, C., Sinigaglia, M. and Spano, G. (2001) Growth inhibition by glucose oxidase system of enterotoxic *Escherichia coli* and *Salmonella derby: in vitro* studies. World J. Microb. Biot. *17*, 287–291.

Matsui, I., Nakashima, K. and Taniguchi, T. (1982) Preparation of cholesterol oxidase. Patent No.: JP57122792.

Matsunaga, I. and Shiro, Y. (2004) Peroxide-utilizing biocatalysts: structural and functional diversity of heme-containing enzymes. Curr. Opin. Chem. Biol. *8*, 127–32.

Medina, M. Vrielink, A. and Cammack, R. (1997) Electron spin echo envelope modulation studies of the semiquinone anion radical of cholesterol oxidase from *Brevibacterium sterolicum*. FEBS Lett. *400*, 247–251.

Nakamatsu, T., Akamatsu. T., Miyajima, R. and Shiio, I. (1975) Microbial production of glucose oxidase. Agric. Biol. Chem. *39*, 1803–1811.

Navas, J., Gonzalez-Zorn, B., Ladron, N., Garrido, P. and Vazquez-Boland, J.A. (2001) Identification and mutagenesis by allelic exchange of *choE*, encoding a cholesterol oxidase from the intracellular pathogen *Rhodococcus equi*. J. Bacteriol. *183*, 4796–4805.

Ohlmeyer, D.W. (1957) Use of glucose oxidase to stabilize beer. Food Technol. October 503–507.

Opwis, K., Knittel, D. and Schollmeyer, E. (2004) Immobilization of catalase on textile carrier materials. AATCC Rev. *4*, 25–28.

Ough, C.S. (1975) Further investigations with glucose oxidase-catalase systems for use with wine. Am. J. Enol. Viticult. *26*, 30–36.

Paar, A., Costa, S., Tzanov, T., Gudelj, M., Robra, K.H., Cavaco-Paulo, A. and Gübitz, G.M. (2001) Thermo-alkali-stable catalases from newly isolated *Bacillus* sp. for the treatment and recycling of textile bleaching effluents. J. Biotechnol. *89*, 147–153.

Paar, A., Raninger, A., de Sousa, F., Beurer, I., Cavaco-Paulo, A. and Gübitz, G.M. (2003) Production of catalase-peroxidase and continuous degradation of hydrogen peroxide by an immobilized alkalothermophilic *Bacillus* sp. Food Technol. Biotech. *41*, 101–104.

Parpinello, G.P., Chinnici, F., Versari, A. and Riponi, C. (2002) Preliminary study on glucose oxidase-catalase enzyme system to control the browning of apple and pear purées. LWT-Food Sci. Technol. *35*, 239–243.

Pazur, J.H. and Kleppe, K. (1964) The oxidation of glucose and related compounds by glucose oxidase from *Aspergillus niger*. Biochemistry *3*, 578–583.

Pena, N., Ruiz, G., Reviejo, A.J. and Pingarron, J.M. (2001) Graphite-teflon composite bienzyme electrodes for the determination of cholesterol in reversed micelles. Application to food samples. Anal. Chem. *73*, 1190–1195.

Petruccioli, M., Ceccarelli, M. and Federici, F. (1993) Screening of *Penicillium* species for the production of glucose oxidase. World J. Microb. Biot. *9*, 77–79.

Pickering, G.J., Heatherbell, D.A. and Barnes, M.F. (1998) Optimising glucose conversion in the production of reduced alcohol wine using glucose oxidase. Food Res. Int. *31*, 685–692.

Pickering, G.J., Heatherbell, D.A. and Barnes, M.F. (1999) The production of reduced-alcohol wine using glucose oxidase treated juice. Part I. Composition. Am. J. Enol. Viticult. *50*, 291–298.

Pickup, J.C., Hussain, F., Evans, N.D., Rolinski, O.J. and Birch D.J.S. (2005) Fluorescence-based glucose sensors. Biosens. Bioelectron. *20*, 2555–2565.

Popper, L. and Knorr, D. (1997) Inactivation of yeast and filamentous fungi by the lactoperoxidase-hydrogen peroxide-thiocyanate system. Nahrung *41*, 29–33.

Primo-Martin, C., Valera, R. and Martinez-Anaya, M.A. (2003) Effect of pentosanase and oxidases on the characteristics of doughs and the glutenin macropolymer (GMP). J. Agric. Food Chem. *51*, 4673–4679.

Primo-Martin, C., Martinez-Anaya, M.A. and Collar, C. (2004) Composition of glutenin macropolymer: effect of flour quality and nonamylolytic enzyme addition. Eur. Food Res. Technol. *218*, 428–436.

Primo-Martin, C., Wang, M.W., Lichtendonk, W.J., Plijter, J.J. and Hamer, R.J. (2005) An explanation for the combined effect of xylanase-glucose oxidase in dough systems. J. Sci. Food Agric. *85*, 1186–1196.

Ram, M.K., Bertoncello, P., Ding, H., Paddeu, S. and Nicolini, C. (2001) Cholesterol biosensors prepared by layer-by-layer technique. Biosens. Bioelectron. *16*, 849–856.

Rasiah, I.A., Sutton, K.H., Low, F.L., Lin, H.M. and Gerrard, J.A. (2005) Crosslinking of wheat dough proteins by glucose oxidase and the resulting effects and bread and croissants. Food Chem. *89*, 325–332.

Revol-Junelles, A.M., Boussouel, N., Ramet, J.P. and Millière, J.B. (2001) Antibacterial activities of lactoperoxidase systems (LPS) modified by I^- and IO_3^- anions. Milchwissenshaft *56*, 329–332.

Reyes De Corcuera J.I., Cavalieri, R.P. and Powers J.R. (2005) Improved platinization conditions produce a 60-fold increase in sensitivity of amperometric biosensors using glucose oxidase immobilized in poly-*o*-phenylenediamine. J. Electroanal. Chem. *575*, 229–241.

Ricci, F. and Palleschi, G. (2005) Sensor and biosensor preparation, optimisation and applications of Prussian Blue modified electrodes. Biosens. Bioelectron. *21*, 389–407.

Ryabov, A.D., Firsova, Y.N, Ershov, A.Yu. and Dementiev, I.A. (1999) Spectrophotometric kinetic study and analytical implications of the glucose oxidase-catalyzed reduction of $[M^{III}(LL)_2Cl_2]^+$ complexes by D-glucose (M = Os and Ru, LL = 2, 2′-bipyridine and 1,10-phenanthroline type ligands). J. Biol. Inorg. Chem. *4*, 175–182.

Sakka, K., Omiya, K. and Shimada, K. (1994) Recombined DNA containing cholesterol oxidase gene and production method using same. Patent Mo.: JP6189754.

Salimi, A., Noorbakhsh, A. and Ghadermarz, M. (2005) Direct electrochemistry and electrocatalytic activity of catalase incorporated onto multiwall carbon nanotubes-modified glassy carbon electrode. Anal. Biochem. *344*, 16–24.

Sampson, N.S., Kass, I.J and Ghoshroy, K.B. (1998) Assessment of the role of an omega loop of cholesterol oxidase: a truncated loop mutant has altered substrate specificity. Biochemistry *37*, 5770–5778.

Sandholm, M., Ali-Vehmas, T., Kaartinen, L. and Junnila, M. (1988) Glucose oxidase (GOX) as a source of hydrogen peroxide for the lactoperoxidase (LPO) system in milk: antibacterial effect of the GOX-LPO system against mastitis pathogens. J. Vet. Med. B *35*, 346–352.

Schumacher, D., Vogel, J. and Lerche, U. (1994) Construction and application of an enzyme electrode for determination of galactose and galactose-containing saccharides. Biosens. Bioelectron. *9*, 85–90.

Singh, S., Chaubey, A. and Malhotra, B.D. (2004) Amperometric cholesterol biosensor based on immobilized ChEt and cholesterol oxidase on conducting polypyrrole films. Anal. Chim. Acta *502*, 229–234.

Situmorang, M., Alexander, P.W. and Hibbert, D.B. (1999) Flow injection potentiometry for enzymatic assay of cholesterol with a tungsten electrode sensor. Talanta *49*, 639–649.

Switala, J., O'Neil, J.O. and Loewen, P.C. (1999) Catalase HPII from Escherichia coli exhibits enhanced resistance to denaturation. Biochemistry *38*, 3895–3901.

Switala J. and Loewen P.C. (2002) Diversity of properties among catalases Arch. Biochem. Biophys. *401*, 145–154.

Takegawa, K., Fujiwara, K., Iwahara, S., Yamamoto, K. and Tochikura, T. (1989) Effect of deglycosylation of *N*-linked sugar chains on glucose oxidase from *Aspergillus niger*. Biochem. Cell Biol. *67*, 460–464.

Takeuchi, K. and Isobe, K. (1999) Production of acidproof catalase. Patent No.: JP11046760.

Tan, X.C., Li, M.J., Cai, P.X., Luo, L.J. and Zou, X.Y. (2005) An amperometric cholesterol biosensor based on multiwalled carbon nanotubes and organically modified sol-gel/chitosan hybrid composite film. Anal. Biochem. *337*, 111–120.

Tarhan, L. (1995) Use of immobilized catalase to remove H_2O_2 used in the sterilisation of milk. Proc. Biochem. *30*, 623–628.

Tzanov, T., Costa, S., Gübitz, G.M. and Cavaco-Paulo, A. (2001) Effect of temperature and bath composition on the dyeing of cotton with catalase-treated bleaching effluent. Color Technol. *117*, 166–170.

Tzanov, T., Costa, S.A., Gübitz, G.M. and Cavaco-Paulo, A. (2002) Hydrogen peroxide generation with immobilized glucose oxidase for textile bleaching. J. Biotechnol. *93*, 87–94.

Vemulapalli, V., Miller, K.A. and Hoseney, R.C. (1998) Glucose oxidase in breadmaking systems. Cereal Chem. *75*, 439–442.

Visser, J. (1991) Biochemical and molecular approaches in understanding carbohydrate metabolism in *Aspergillus niger*. J. Chem. Tech. Biot. *50*, 111–113.

Volotovsky, V. and Kim, N. (1998) Determination of glucose, ascorbic and citric acids by two-ISFET multienzyme sensor. Sensor. Actuat. B-Chem. *49*, 253–257.

Wang, J. (1999) Amperometric biosensors for clinical and therapeutic drug monitoring: a review. J. Pharmaceut. Biomed. *19*, 47–53.

Wang, J., Lu, F., Angnes, L., Liu, J., Sakslund, H., Chen, Q., Pedrero, M., Chen, L. and Hammbich, O. (1995) Remarkably selective metallized-carbon amperometric biosensors. Anal. Chim. Acta *305*, 3–7.

Watanabe, K., Aihara, H. and Nakamura, M. (1986) Production of cholesterol oxidase. Patent No.: JP61247381.

Whittaker, M.M. and Whittaker, J.W. (1988) The active site of galactose oxidase. J. Biol. Chem. *263*, 6074–6080.

Whittaker, M.M. and Whittaker, J.W. (1990) A tyrosine-derived free radical in apogalactose oxidase. J. Biol. Chem. *265*, 9610–9613.

Whittaker, M.M. and Whittaker, J.W. (2000) Expression of recombinant galactose oxidase by *Pichia pastoris*. Protein Expres. Purif. *20*, 105–111.

Wilson, G.S. and Gifford, R. (2005) Biosensors for real-time *in vivo* measurements. Biosens. Bioelectron. *20*, 2388–2403.

Wohlfahrt, G., Witt, S., Hendle, J., Schomburg, D., Kalisz, H.M. and Hecht, H.-J. (1999) 1.8 and 1.9 Å resolution structures of the *Penicillium amagasakiense* and *Aspergillus niger* glucose oxidases as a basis for modelling substrate complexes. Acta Crystallogr. D Biol. Crystallogr. *55*, 969–977.

Wu, B.L., Zhang, G.M., Shuang, S.M., Dong, C., Choi, M.M.F. and Lee, A.W.M. (2005) A biosensor with myrosinase and glucose oxidase bienzyme system for determination of glucosinolates in seeds of commonly consumed vegetables. Sensor. Actuat. B-Chem. *106*, 700–707.

Wu, A.J., Penner-Hahn, J.E. and Pecoraro, V.L. (2004) Structural, spectroscopic, and reactivity models for the manganese catalases. Chem. Rev. *104*, 903–938.

Wu, X.J. and Choi, M.M.F. (2003) Hydrogel network entrapping cholesterol oxidase and octadecyl-silica for optical biosensing in hydrophobic organic or aqueous micelle solvents. Anal. Chem. *75*, 4019–4027.

Xie, L.L. and van der Donk, W.A. (2001) Homemade cofactors: self-processing in galactose oxidase. Proc. Natl. Acad. Sci. USA *98*, 12863–12865.

Xu, F., Golightly, E.J., Schneider, P., Berka, R.M., Brown, K.M., Johnstone, J.A., Baker, D.H., Fuglsang, C.C., Brown, S.H. and Klotz, A.V. (2000) Expression and characterisation of a recombinant *Fusarium spp*. galactose oxidase. Appl. Biochem. Biotech. *88*, 23–32.

Yu, J.H., Liu, S.Q. and Ju, H.X. (2003) Glucose sensor for flow injection analysis of serum glucose based on immobilization of glucose oxidase in titania sol-gel membrane Biosens. Bioelectron. *19*, 401–409.

CHAPTER 26

LACCASES: BIOLOGICAL FUNCTIONS, MOLECULAR STRUCTURE AND INDUSTRIAL APPLICATIONS

MIGUEL ALCALDE*

Departamento de Biocatálisis, Instituto de Catálisis y Petroleoquímica, CSIC, Cantoblanco, Madrid, Spain

**malcalde@icp.csic.es*

1. INTRODUCTION

Laccases (benzenediol:oxygen oxidoreductase, EC 1.10.3.2) are extracellular, multicopper enzymes that use molecular oxygen to oxidize various aromatic and non-aromatic compounds by a radical-catalysed reaction mechanism (Thurston, 1994). They belong to a larger group of enzymes termed the blue-multicopper oxidase family which includes the plant ascorbate oxidase, the mammalian plasma protein ceruloplasmin and bilirubin oxidase, among others. The term Laccase stems from its original identification in the exudates of the Japanese lacquer tree *Rhus vernicifera* described by Yoshida (1883). Just over a century later it was characterized as being a metal-containing oxidase by Bertrand (1985). Laccases have also been found in other plants, and also insects and bacteria, but are predominant in fungi. Laccase activity has been demonstrated in more than 60 fungal strains belonging to Basidiomycetes, Ascomycetes and Deuteromycetes, being documented in virtually every fungus examined for it (Gianfreda *et al.*, 1999). Its presence in plants appears to be far more limited than in fungi. All species of family Anacardiaceae, of which the lacquer tree is member, contain laccase in the resin products and in the secreted resin (Huttermann *et al.*, 2001). Reports on the presence of laccase in other plants are however limited to *Acer pseudoplatanus*, *Pinus taeda*, *Aesculus parviflora* and *Populus eruamericana* (Mayer and Staples, 2002), though it is believed that they are present throughout the plant kingdom. Polyphenol oxidases, perhaps laccase-like, have also been reported in insects (Parkinson *et al.*, 2003), and there is strong evidence for the widespread distribution of laccases in prokaryotes. The first bacterial laccase to be extensively studied was from *Axospirillum lipoferum* and the crystal structure of a

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 461–476

bacterial (*Bacillus subtilis*) laccase is now available (Enguita *et al.*, 2003). Sequence homology analysis suggests that laccases also occur in bacteria such as *Mycobacterium tuberculosis* (Alexandre and Zhulin, 2000). Laccase encoding genes have been found in Gram-negative and Gram-positive bacteria, including species living in extreme habitats, *e.g.* in *Oceanobacillus iheyensis* and *Aquifex aeolicus* and in the archaeobacterium *Pyrobaculum aerophilum* (Claus, 2004, 2003). The development of metagenomics in the next few years can be expected to lead to the discovery of novel bacterial laccases from extremophile environments.

2. LACCASE SUBSTRATES AND INHIBITORS

Laccase catalyses the four electron reduction of molecular oxygen to water with one-electron oxidation of reducing substrate, without producing hydrogen peroxide. Although laccases preferably act on phenolic compounds their substrate spectrum is huge, and since the range of substrates oxidized varies from one laccase to another this complicates the formulation of a precise definition of laccase activity. In addition there is overlap in substrate range with another type of (copper-containing) oxidase – tyrosinase -, notionally a mono-phenol monooxygenase. Laccases can convert *o*- and *p*-diphenols, aminophenols, methoxy-substituted phenols, benzenethiols, polyphenols, polyamines, hydroxyindols, some aryl diamines and a considerable range of other compounds but do not oxidize tyrosine (whereas the tyrosinases do). Inorganic/organic metal compounds are also substrates of laccase. Mn^{2+} is oxidized to Mn^{3+} and also $Fe(EDTA)^{2-}$ is accepted by the enzyme. All known laccases catalyse the oxidation of ascorbic acid and phenol substrates with equally high efficiencies. Simple diphenols like hydroquinone and catechol are generally good substrates but guaicol and 2,6-dimethoxyphenol are often better. *p*-Phenylene diamine is a frequently used substrate. Syringaldizine (N, N′-bis(3,5-dimethoxy-4-hydroxybenzylidene hydrazine; $\varepsilon_{525} = 65000\,M^{-1}cm^{-1}$) is a good substrate but it has to be used in the complete absence of hydrogen peroxide since syringaldizine is also oxidized by the manganese-dependent peroxidases produced by many lignolytic basidiomycetes (Thurston, 1994; Gianfreda *et al.*, 1999.)

Laccases have wide reducing substrate (DH) specificity - the K_m^{DH} is in the range of 1–10 mM -. By contrast, the enzyme displays a strong preference for O_2 as the oxidizing substrate. K_m^{O2} values are around 10^{-5} M and V_{max} depends on the source of the laccase ($50–300\,Msg^{-1}$).

Laccases can be very strongly inhibited by various reagents. Small anions such as azide, halides, cyanide, thiocyanide, fluoride and hydroxide bind to the type 2 and type 3 Cu, resulting in the interruption of internal electron transfer and inhibition of activity. Other inhibitors include metal ions (*e.g.* Hg^{2+}), fatty acids, sulfhydryl reagents, hydroxyglycine, kojic acid, desferal and cationic quaternary ammonium detergents the reactions with which may involve amino acid residue modifications, conformational changes or Cu chelation (Gianfreda *et al.*, 1999; Call and Mucke, 1997). Regarding conformational changes, it is known that these are highly depended on the state of oxidation of the copper atoms. This is one of

the reasons for the sensitivity towards chelating agents. The selective removal of Cu by chelating agents (EDTA, dimethyl glyoxime, N, N′-diethyldithiocarbamate, NTA) leads to a loss of catalytic activity. Laccases generally are more stable at alkaline pH than at acidic pH, probably due to the OH^- inhibition of auto-oxidation.

3. STRUCTURAL AND FUNCTIONAL FEATURES

Phylogenetic analyses based on sequence alignments have shown that laccase copper-binding domains are highly conserved even when the rest of the molecule shows wide variability. The topology of the phylogenetic trees obtained suggests that a single monophyletic branch might exist for the fungal laccases (Valderrama *et al*., 2003). It is assumed that laccases are evolutionarily very ancient enzymes, and that the enzyme activity linked to three different copper sites must have arisen early in the course of evolution. In a recent study, sequence alignments of more than 100 plant and fungal laccases resulted in the identification of a sequence signature that uniquely characterizes the laccases as a distinctive subgroup of enzymes of the multi-copper family. This signature, comprising 4 ungapped sequence segments L1-L4 ranging from 8 to 24 residues in length and scattered across almost the entire length of the protein, contains the twelve amino acid residues that serve as the copper ligands (Kumar *et al*., 2003). This structural conservation reflects a common reaction mechanism for copper oxidation and O_2 reduction in these enzymes.

3.1. Metal Content

Laccases contain four copper atoms per monomer and no other co-factor (Davies, 2002). The four copper atoms are bound to 3 redox sites designated T1, T2 and T3. The four complexed copper atoms are designated Type I, Type II and Type III, according to their spectroscopic and paramagnetic properties.

Paramagnetic “blue” Cu1 (copper Type I, T1) is bound as a mononuclear cluster and confers the beautiful greenish-blue colour to laccases (in the oxidized resting state, $\varepsilon_{610nm} \sim$ is in the range 4900–5700 $M^{-1}cm^{-1}$). This intense absorption band results from the covalent copper-cysteine bond. In fact the Cys $\rightarrow Cu^{2+}$ charge transfer yields an electron paramagnetic resonance (EPR) spectrum with an exceptionally small hyperfine splitting in the low-field region. Cu T1 is available to solvents, including water. It can be removed from the enzyme molecule by various copper complexones, or substituted by cobalt. It can also be displaced by mercury but this affects severely enzyme activity.

Paramagnetic “normal or non-blue” Cu2 (copper Type II, T2) behaves as a mononuclear site with normal EPR spectrum parameters attesting the presence of tetragonal surroundings. It has only weak and insignificant absorption in the UV-VIS region but is EPR-active.

Diamagnetic spin-coupled Cu3–Cu4 pair (coppers Type III, T3) form a binuclear site in which the two copper atoms are anti-ferromagnetically coupled (the so called Cu3–Cu4 dyad) through a bridging ligand – an hydroxyl bridge-. The main consequence is the total absence of an EPR signal. The T3 pair are most likely responsible for a shoulder in the UV spectrum at about 330nm (oxidized form) which disappears upon reduction of the active site. Together the T2 and T3 coppers form a trinuclear cluster where reduction of molecular oxygen and the release of water takes place. The T1 copper is involved in the oxidation of the reducing substrate, capturing electrons that are then transferred to the T2 and T3 copper centres.

3.2. Redox potential, E°

Redox potential (E°) *i.e.* the energy required to capture one electron from a reducing substrate with the corresponding formation of a cation radical, is one of the most significant features of laccases. The E° of laccases is typically determined using an appropriate couple such as $K_3Fe(CN)_6 - K_4Fe(CN)_6$ (+433 mV). Despite the strong similarities in their EPR parameters, the redox potentials of T1 centres can vary widely between laccases from different sources. For example, the E° of fungal laccases is far higher than that of plant or bacterial laccases. Thus, the *Rhus vernicifera* laccase has a Cu T1 E° of about +400 mV, whereas the E° of fungal laccases range from +400 to +800 mV. Ligninolytic peroxidases are able to oxidize substrates of extremely high electropotential, up to 1.49V (*e.g.* lignin peroxidase can convert veratryl alcohol whose E° = 1.22 V). Laccases, however, cannot oxidize non-phenolic substrates with electropotentials greater than 1.06 V (Table I).

The reactivity of laccases has been correlated with their redox potential which is thought to play a major role in the overall performance of these enzymes (Xu *et al.*, 1998). It has been demonstrated that the oxidation rate depends on the E° difference between the reducing substrate and the Cu T1. A lower E° of substrate or a higher E° of laccase (Cu T1) often results in a higher rate of substrate oxidation. The broad difference observed between the E° values of different laccases (from +465 mV for *Myceliophthora thermophyla* laccase to +780 mV of *Trametes – Polyporus* or *Coriolus – versicolor*) has been the subject of studies aimed to determine which parameters modulate E°. It is well known that a hydrophobic residue (either Phe or Leu) at the axial position of the T1 Cu site is implicated in the elevated redox potential of fungal laccases (see Table I). However, there are broad differences in redox potentials even amongst the fungal laccases; thus it is clear that a Leu or Phe residue at the T1 site cannot be the sole contributor to these effects. In this sense, the significance of a highly conserved pentapeptide segment located near the Cu T1 has been described (Xu *et al.*, 1998). Recently, analysis of sequence alignments, site-directed mutagenesis experiments and crystallographic studies have allowed to propose mechanisms that explain how laccases can tune their redox potential by as much as 200 mV (Piontek *et al.*, 2002).

Table 1. E° of several blue multicopper enzymes.

Species	Organism	Enzyme	CuT1**E**°(V)	Potential axial ligand	Ref
Trametes versicolor	Basidiomycete	Laccase	+0.79	Phe	(Alcalde *et al.*, 2002)
Trametes villosa	Basidiomycete	Laccase	+0.79	Phe	(Kumar *et al.*, 2003)
Neurospora crassa	Ascomycete	Laccase	+0.78	Leu	(Piontek *et al.*, 2002)
Rhizoctonia solani	Deuteromycete	Laccase	+0.71	Leu	(Kumar *et al.*, 2003)
Coprinus cinereus	Basidiomycete	Laccase	+0.55	Leu	(Kumar *et al.*, 2003)
Scytalidium thermophilum	Basidiomycete	Laccase	+0.51	Leu	(Kumar *et al.*, 2003)
Homo sapiens	Mammalian	Ceruloplasmin	+0.49	Met	(Kumar *et al.*, 2003)
Myrothecium verrucaria	Fungi mitosporic	Bilirubin oxidase	+0.48	Met	(Kumar *et al.*, 2003)
Myceliophthora thermophila	Ascomycete	Laccase	+0.47	Leu	(Alcalde *et al.*, 2002)
Rhus vernicifera	Plant	Laccase	+0.43	Met	(Yaropolov *et al.*, 1994)
Zucchini (Cuburbita pepo)	Plant	Ascorbate oxidase	+0.34	Met	(Kumar *et al.*, 2003)

3.3. Protein Structure

The crystal structures of one bacterial laccase (the spore coat protein from *Bacillus subtilis*, BsL) (Enguita *et al.*, 2003) and three fungal laccases (*Coprinus cinereus* (CcL) Ducros *et al.*, 1998; *Trametes versicolor* (TvL) Piontek *et al.*, 2002; *Melanocarpus albomyces* (MaL) Hakulinen *et al.*, 2002) have been reported. All were resolved at high resolutions: 1.7; 2.2, 1.90 and 2.4 Å, respectively. In the cases of MaL and TvL, the structures of the active enzymes containing a full complement of copper atoms were determined. Deglycosylation carried out to obtain diffracting crystals of CcL resulted in the loss of copper (CuT2), therefore the available structure corresponds to a catalytically incompetent state.

Fig. 1 shows the overall structure of laccase from *M. albomyces* (MaL). The structures of MaL and TvL are similar. Both are monomers consisting of three cupredoxin-like domains. One of these domains (domain 3) contains the mononuclear site T1. The trinuclear cluster (T2 and T3) is embedded between domains 1 and 3 both of which provide residues for the coordination of the coppers. Domain 2 contains residues that participate in substrate binding. MaL and TvL are

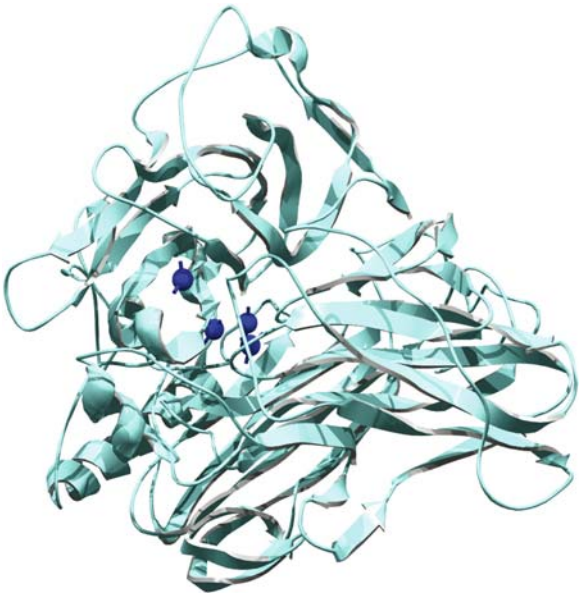


Figure 1. Overall structure of *Melanocarpus albomyces* laccase (MaL) according to Hakulinen *et al.* (2002). Copper atoms are drawn as dark spheres. Disulfide bonds are included as stick models

both heavily glycosylated (between 5–9 glycosylation sites) and their structures are stabilized by two (TvL) or three (MaL) disulfide bridges.

Fig. 2 shows the disposition of the copper atoms in MaL. The mononuclear site contains one Cu T1, which is trigonally coordinated to two ND atoms from two His residues and a SG atom from a Cys. Additionally, the T1 copper is connected to the trinuclear cluster by a His-Cys-His tripeptide which is highly conserved among the blue multicopper oxidases. The closest distance between the T1 and T2/T3 coppers is about 12 Å. The type-1 centres of the blue multicopper family usually have an additional axial ligand provided by a methionine sulphur atom (*i.e.* ascorbate oxidase has a SD atom from a Met residue 2.9 Å away from Cu T1, see Table 1). Fungal laccases, however, have a Phe or Leu residue at the equivalent position (Phe for TvL and Leu for MaL). In the TvL and MaL structures, Phe and Leu are 3.6 and 3.7 Å from the Cu T1, respectively, and do not participate in the coordination. Furthermore, no residue occupies the axial position on the other side, so this position is free for the substrate. Thus, the coordination – trigonal coplanar - of the T1 copper in TvL or MaL (and in general in all fungal laccases) is different from that of the rest of the blue multicopper family which consists of two histidines, one cysteine and one axial methionine and is therefore 4-fold.

Axial coordination has been considered to be one factor affecting the redox potential of copper enzymes. The geometry of the T2/T3 cluster is very similar to that found in the crystal structure of ascorbate oxidase. The three coppers are

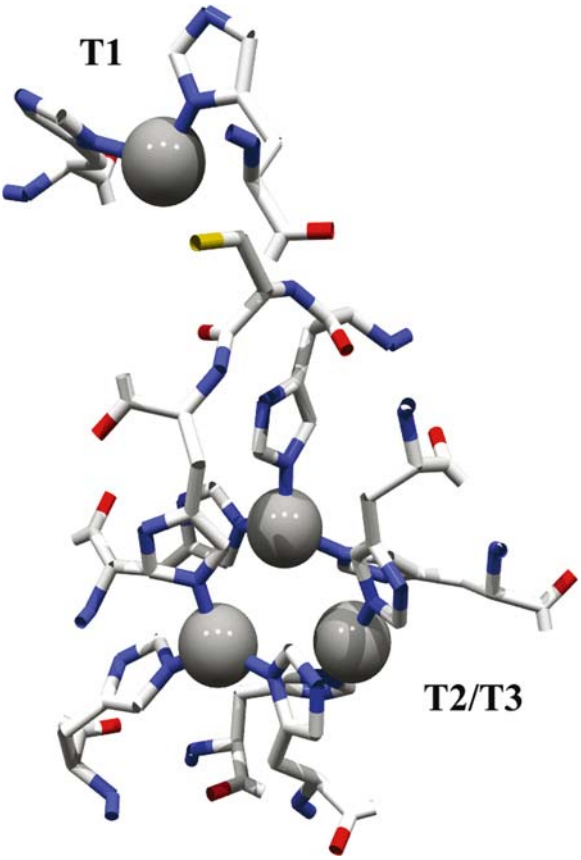


Figure 2. Geometry of the four copper atoms at the T1 and T2/T3 sites in the *Melanocarpus albomyces* laccase structure (Hakulinen *et al.*, 2002)

arranged in an almost perfect regular triangle. In this site two T3 coppers (Cu3 and Cu4) are coordinated symmetrically to six His N atoms. The ligation of each of these coppers is therefore 4-fold, and their coordination sphere can be best described as being a distorted tetrahedron. Cu T2 is coordinated to two His N atoms and to one atom that is probably a chloride ion. Cu T2 is more exposed and more labile compared with the other two coppers at the T3 site. In this regard it is the T2 copper site that is deficient in copper in the copper-depleted forms of both laccase and ascorbate oxidase. Two channels, one broad and one narrow, facilitate access of solvent molecules to the trinuclear site in ascorbate oxidase and also in MaL, CcL and TvL. The narrow channel leads to the T2 copper. The broad channel, which is approximately 10 Å long and leads to one of the type-3 coppers (Cu3), is located between the trinuclear centre and the protein surface. It is worth noting that the crystal structures reveal one exclusive feature in MaL that TvL does not have: a

plug formed by the last four C-terminal residues of the polypeptide. This plug is a consequence of the C-terminal processing common to other ascomycete laccases (Kiiskinen and Saloheimo, 2004; Bulter *et al.*, 2003). From crystallographic and phylogenetic studies one can conclude that the active site and its environment are structurally highly conserved. This is true for the copper geometry, for the two channels which provide access for molecular oxygen and for the conserved His-Cys-His tripeptide implicated in the electron transfer pathway between the T1 copper and the trinuclear cluster. This structural conservation reflects a common mechanism for copper oxidation and O_2 reduction in these enzymes.

3.4. Catalytic Mechanism

Neither the electron transfer mechanism nor the oxygen reduction to water are fully understood. However several facts are well established:

1) Laccases react oxidizing the reducing substrate by T1 Cu^{2+}-mediated abstraction of one electron. Consequently a free (cationic) radical is formed. This radical can further undergo laccase-catalysed oxidation (*e.g.* phenol to quinone) or non-enzymatic reactions (*e.g.* hydration or polymerisation).
2) As a one-electron substrate oxidation is coupled to the four-electron reduction of oxygen the reaction mechanism cannot be entirely straightforward. Laccase can be thought to operate as a battery, storing electrons from individual oxidation reactions in order to reduce molecular oxygen. Hence the oxidation of four reducing substrate molecules is necessary for the complete reduction of molecular oxygen to water.
3) Each electron extracted from the four monoelectronic oxidations at the T1 site is transferred to the trinuclear cluster where O_2 is bound. Thus the T2 and T3 sites are the locations where reduction of molecular oxygen and the release of water occurs.

A “two-site Ping-Pong Bi-Bi” reaction mechanism has been established for laccase, which means that products are released before the binding of new substrate occurs. It appears that the solvent channels of the blue copper oxidases are well suited to allow fast access of dioxygen molecules to the trinuclear cluster and subsequent easy release of water. Although many catalytic schemes have been proposed for this (Torres *et al.*, 2003; Davies, 2002) the major unknown is the reductive part of the cycle: the mechanism by which the trinuclear cluster is reduced.

4. BIOLOGICAL FUNCTIONS

Different physiological roles have been suggested for laccases, including plant wounding response, the development of fruiting bodies, cell-wall reconstitution and metabolic turnover of soil humic matter. Laccases are also believed to play roles in pathogenesis (fungal virulence factors), sporulation, fungal spore pigmentation and

fungal morphogenesis in general. One of the most controversial and more studied biological functions of laccases is related to the process of lignification of plant cell walls and lignin polymerisation during white rot of wood.

Lignin, which is a structural component of the plant cell wall, is a heterogeneous and complex polyphenolic biopolymer that consists of phenyl propanoid units linked by various non-hydrolysable C-C and C-O bonds. The considerable abundance of wood-rotting Basidiomycete fungi as laccase producers seems to indicate that the main role of fungal laccase is to depolymerise lignin. However, this function contrasts with that of laccases in plants which are components of the lignin-synthesizing system. Whereas enzymes in general are highly substrate specific, a remarkable property of lignolytic enzymes is the breadth of their substrate range. The most important lignin degrading enzymes are lignin peroxidases, manganases peroxidases and laccases but some other enzymes such as celobiose:quinone oxidoreductase, cellobiose dehydrogenase, glyoxalate oxidase, glucose oxidases (glucose 1-oxidase and pyranose 2-oxidase), veratryl alcohol oxidase and some esterases may also play roles in the complex process of natural wood decay. Consequently, the lignin biodegradation process involves the synergistic effects of many enzymes and non-enzymatic components (mediators) that interact to reach an equilibrium between enzymatic polymerisation and depolymerisation. Indeed, some experimental evidence suggests that laccase acting on lignin may display both activities. What is clear is that the enzymes assumedly involved in lignin cleavage produce highly reactive (and hence highly toxic) species from which the fungal mycelium must be protected. It might be that one of the functions of laccase is to scavenge these compounds by promoting polymerisation before they can enter the hypha. At the present however, the information in hand is insufficient for definite conclusions to be reached on the role of laccases in the lignification/delignification process.

5. INDUSTRIAL APPLICATIONS

Laccases are increasingly being used in a wide variety of industrial oxidative processes such as delignification, dye or stain bleaching, bioremediation, plant-fibre modification, ethanol production, biosensors, biofuel cells *etc*. Industrial uses require overproduction of the enzyme, generally in a heterologous host, as an indispensable prerequisite. Indeed, most commercial laccases are produced in *Aspergillus* hosts. The functional expression of the *Myceliophthora thermophila* laccase in *S. cerevisiae* by directed molecular evolution has been reported, which enables this system to be tuned up for new and challenging applications (Bulter *et al*., 2003). In recent work an efficient transformation and expression system was developed for the basidiomycete *Pycnoporus cinnabarinus* and this was used to transform a laccase-deficient monokaryotic strain with the homologous lac1 laccase gene. The yield obtained was as high as 1.2 g of laccase per litre and represents the highest laccase production reported for recombinant fungal strains (Alves *et al*., 2004).

5.1. The Laccase-mediator System (LMS)

Over the last 10 years the versatility of laccase has been broadened by the invention of the laccase-mediator system (LMS). The combination of the enzyme with low molecular weight molecules such as 2, 2′-azinobis (3-ethylbenzthiazoline-6-sulfonate) (ABTS) or 1-hydroxybenzotriazole (HBT) not only leads to higher rates in the conversion of known substrates but also adds new reactions for which laccase alone had no or only marginal activity. The study of the LMS in the pulp-kraft bleaching industry or in the removal of harmful xenobiotics (*e.g.* polycyclic aromatic hydrocarbons, PAHs) is well documented (Call and Mucke, 1997). Generally, all mediators are substrates of laccases (*e.g.* syringaldazine). They are easily oxidized at the T1 site, in some cases producing very unstable and reactive cationic radicals which can in turn oxidize more complex substrates. In this mechanism, the mediator acts as a diffusible electron carrier (Majcherczyk and Johannes, 2000) permitting higher molecular weight substrates such as lignin can be oxidised. The electrons acquired by the laccase molecule are finally transferred to oxygen to form water (at the T2/T3 trinuclear cluster). The action of a mediator is advantageous as it enables laccases to achieve two goals: i) the oxidation of polymers by side-stepping the inherent steric hindrance problems (enzyme and polymer do not have to interact in a direct manner) and ii) increased substrate range, being able to oxidize compounds having redox potentials exceeding their own. An effective mediator does not necessarily have to have a redox potential higher than that of the laccase. A potential higher than that of the substrate and free diffusiblity are generally more important. It is also worth pointing out that the synergistic effects of mixtures of mediators can improve the oxidation by laccases (Pickard *et al.*, 1999). However, chemical mediators are mostly toxic, unstable or expensive. Moreover, they lead to by-products and inactivate the enzyme: by oxidizing the mediator, laccase is generating a strongly oxidizing intermediate the "co-mediator" which apart from acting as a diffusible electron carrier also interacts with the laccase inactivating the biocatalyst. Novel approaches to overcome this shortcoming are currently being developed (from searching for natural mediators *e.g.* tyrosine (Johannes and Majcherczyk, 2000; Camarero *et al.*, 2005) to the directed evolution of laccases (Bulter *et al.*, 2003).

5.2. LMS in the Pulp and Paper Industry

The removal of lignin from woody tissues is a process that has attracted a very great deal of research, especially due to its importance in the pulp and paper industry. Bleaching of pulp for paper manufacture is carried out in some instance using harmful and polluting chemicals. Traditional techniques utilize chlorine-based agents – including Cl_2 - that can lead to the release of toxic contaminants to the environment. Bourbonnais and Paice (1990) demonstrated for the first time that a laccase from *T. versicolor* efficiently demethylated and delignified kraft pulp. Since then a number of studies have been carried out analysing most of the factors

involved in the process (Call and Call, 2005; Sigoillot *et al.*, 2005; Bajpai, 2004; Call and Mucke, 1997). The application of laccases in pulp-kraft bleaching may result in higher pulp yields and energy savings.

5.3. LMS and Polycyclic Aromatic Hydrocarbons

Polycyclic aromatic hydrocarbons (PAHs) are a class of highly dangerous xenobiotics (mutagenic, carcinogenic) widely distributed in terrestrial and aquatic environments. Laccase is able to oxidize both low molecular weight PAHs and the more recalcitrant ones (*i.e.* benzo[*a*]pyrene) to more benign, less toxic compounds. To exert this remarkable biotransformation, laccase needs the presence of redox mediators (basically HBT and ABTS) (Alcalde *et al.*, 2002; Majcherczyk and Johannes, 2000; Majcherczyk *et al.*, 1998; Collins *et al.*, 1996; Johannes *et al.*, 1996). Natural mediators (tyrosine, 4-hydroxybenzoic acid, hydroxybenzilic alcohol) secreted by white-rot fungi have also been demonstrated to be useful for PAH oxidation to quinones (Johannes and Majcherczyk, 2000). It has been reported that the biodegradation of PAHs and their metabolites increases with their oxidation state. This suggests that enzymatic oxidation followed by the action of microorganisms could be an effective PAH remediation strategy. Additionally, quinones are significantly less mutagenic and carcinogenic than their corresponding PAHs (Torres *et al.*, 2003).

5.4. Use of Laccases in Organic Solvents

Due to the hydrophobicity and low aqueous solubility of many laccase substrates (*e.g.* lignin, PAHs, steroids), reactions are usually performed in the presence of organic solvents (Torres *et al.*, 2003; Gianfreda *et al.*, 1999; Yaropolov *et al.*, 1994). However, under these conditions laccases are fairly unstable being either denaturated or inhibited (Rodakiewicz-Nowak *et al.*, 2000; Luterek *et al.*, 1998). In this regard strategies have been reported for improving the performance of laccases in organic solvents *e.g.* enzyme immobilization (Duran *et al.*, 2002) and directed molecular evolution of laccases (Alcalde *et al.*, 2005). In this latter approach the *Myceliophthora thermophila* laccase expressed in *S. cerevisiae* was engineered by *in vitro* evolution in the presence of increasing concentrations of acetonitrile and ethanol. Screening was focused towards higher activity and stability. The turnover rates of mutant enzymes at high concentrations of organic solvents were several fold improved.

5.5. Enzymatic Bioremediation with Laccases

In addition to their effects on PAHs, laccases can diminish the toxicity of phenols, trichlorophenols, organophosphorus pesticides and azo dyes, among others. Phenols and their derivatives (chlorophenols, dimethoxyphenols and nitrophenols) are among the most common organic chemicals found in industrial wastewaters and in sanitary

waste sites. They can be oxidized by laccase (Ullah *et al.*, 2000) resulting in the generation of o-benzoquinones which are less toxic than phenols. Organophosphorus insecticides and nerve agents that contain the P-S bond are highly resistant to enzymatic hydrolysis (Jauregui *et al.*, 2003). Laccases from different sources can perform their oxidation in the presence of chemical mediators (ABTS) (Amitai *et al.*, 1998). An ample range of azo dyes are oxidized by laccases in the presence of mediators. The effect of HBT and ABTS has been extensively studied (Almansa *et al.*, 2004). The recent finding of natural mediators for dye bleaching open new perspectives in the application of the LMS (Camarero *et al.*, 2005).

5.6. Mediator-less Electroreduction of Oxygen to Water

Laccase is one of the few enzymes able to catalyse enzymatic and electrochemical reactions even in the absence of a low molecular weight electron carrier (Call and Mucke, 1997; Yaropolov *et al.*, 1994). A cathode can be used to substitute one of the substrates of the enzymatic reaction as an electron donor. Electrons are therefore taken by the Cu T1 of the enzyme directly from the electrode. The reaction is called mediator-less electroreduction which runs without the formation of hydrogen peroxide as an intermediate. The electrode potential established as a result of the reaction is close to the equilibrium oxygen potential. The overall reaction is the electroreduction of molecular oxygen to water. The most effective inorganic catalyst for this reaction is a specially treated form of platinum (equilibrium oxygen potential: 1.23 V). In contrast, laccase immobilized on the surface of electrodes of various materials is able to change the potentials in the range of 1.2–0.6V depending on the amount of enzyme fixed. It should be noted that the *Trametes versicolor* laccase does not catalyse water decomposition at potentials greater than 1.2 V. This is due to the irreversible inactivation of the laccase at higher potentials by overoxidation of catalytically important groups of the enzyme. A number of potentiometric biosensors based on this principle have been developed (Freire *et al.*, 2003; Freire *et al.*, 2002; Milligan and Ghindilis, 2002). Another significant application is in the engineering of biofuel cells (Shleev *et al.*, 2005). One of the most brilliant works done in this field reported a miniature biofuel cell made from two 7μm-diameter, 2cm long carbon fibre electrodes, glucose oxidase being coupled to the anode and laccase immobilized on the cathode (Chen *et al.*, 2001).

5.7. Medium Density Fibreboards (MDF)

Laccases are being used in the enzymatic cross-linking of lignin-based materials to produce medium density fibreboards (MDF) which are nowadays employed in the construction of furniture. Formerly the manufacture of MDF was based on non-enzymatic cross-linking by harmful and polluting compounds such as formaldehyde (Widsten *et al.*, 2004; Felby *et al.*, 2002).

5.8. Ethanol Production

Dilute acid hydroysis of lignocellulose yields sugar mixtures which can be fermented by yeast to produce fuel alcohol. However, phenolic compounds present in the hydrolysate inhibit yeast fermentation. Heterologous expression of the laccase from *Trametes versicolor* in *S. cerevisiae* during fermentation has recently been demonstrated to improve the production of fuel ethanol from this renewable raw material by eliminating these compounds (Larsson *et al.*, 2001).

5.9. Cosmetics

Bleaching and/or dyeing of hair usually involves the use of harsh chemicals which can damage hair. Dye precursors can be oxidized to the effective colouring agent using laccase instead of the chemical agent. Laccases have also been used in the presence of hydroxy stilbenes as hair bleaches (Onuki *et al.*, 2000; Pruche *et al.*, 2000). Pigments used in cosmetic applications as well as melanins could be produced using laccase catalysed oxidative reactions. Conversely, laccase may also find utility in the reduction of melanin-derived blemishes in topical skin applications (Tetsch *et al.*, 2005; Nagai *et al.*, 2003).

5.10. Food Industry

Laccase can be used in beverage (wine, fruit juice and beer) processing, ascorbic acid determination, sugar beet pectin gelation, baking, as a biosensor and to improve food sensory parameters (Minussi *et al.*, 2002).

Wine stabilization is one of the main applications of laccase in the food industry. Laccase is used to selectively target specific polyphenols during the madeirization process. The use of the laccase in must results in a stable wine with a good flavour. Since the use of laccase as a food additive is still not permitted, this enzyme has been used in wine production in an immobilized form to ensure its elimination from the must and facilitate its reutilization (Servili *et al.*, 2000).

The tendency for hazes to develop in some beers during long-term storage is a persistent problem in the brewing industry. As an alternative to the traditional treatment, laccase can be added to the wort or at the end of the process. Since oxygen is not desirable in the finished beer, addition of laccase can remove any excess oxygen whereby the storage life of beer is enhanced. At the same time the laccase removes some of the polyphenols that may still remain in the beer. The polyphenol complexes formed by laccase may be removed by filtration or other means of separation (Mathiasen, 1995).

For apple and grape juices excessive oxidation of phenolics has almost always been considered detrimental to the organoleptic quality of the product. These beverages are typically stabilized to delay the onset of protein-polyphenol haze formation. Various enzymatic treatment have been proposed for fruit juice stabilization, among them the use of laccase (Minussi *et al.*, 2002).

REFERENCES

Alcalde, M., Bulter, T. and Arnold, F.H. (2002) Colorimetric assays for biodegradation of polycyclic aromatic hydrocarbons by fungal laccases. J. Biomol Screening *7*, 547–553

Alcalde, M., Bulter, T., Zumárraga, M., García-Arellano, H., Mencía, M., Plou, F.J. and Ballesteros, A. (2005) Screening mutant libraries of fungal laccases in the presence of organic solvents. J. Biomol. Screening. *10*, 624–631.

Alexandre, G. and Zhulin, I.B. (2000) Laccases are widspread in bacteria. Trends Biotechnol. *18*, 41–42

Almansa, E., Kandelbauer, A., Pereira, L., Cavaco-Paulo, A. and Guebitz, G.M. (2004) Influence of structure on dye degradation with laccase mediator systems. Biocatalysis and Biotransformation *22*, 315–324

Alves, A.M.C.R., Record, E., Lomascolo, A., Scholtmeijer, K., Asther, M., Wessels, J.G.H. and Wosten, H.A.B. (2004) Highly efficient production of laccase by the basidiomycete *Pycnoporus cinnabarinus*. App. Environ. Microbiol. *70*, 6379–6384

Amitai, G., Adani, R., Sod-Moriah, G., Rabinovitz, I., Vincze, A., Leader, H., Chefetz, B., Leibovitz-Persky, L., Friesem, D. and Hadar, Y. (1998) Oxidative biodegradation of phosphorothiolates by fungal laccase. Febs Letters *438*, 195–200

Bajpai, P. (2004) Biological bleaching of chemical pulps. Crti. Rev. Biotechnol. *24*, 1–58

Bourbonnais, R. and Paice, M.G. (1990) Oxidation of nonphenolic substrates - an expanded role for laccase in lignin biodegradation. Febs Letters *267*, 99–102

Bulter, T., Alcalde, M., Sieber, V., Meinhold, P., Schlachtbauer, C. and Arnold, F.H. (2003a) Functional expression of a fungal laccase in *Saccharomyces cerevisiae* by directed evolution. Appl. Environ. Microbiol. *69*, 987–995

Call, H.P. and Call, S. (2005) New generation of enzymatic delignification and bleaching. Pulp and Paper-Canada *106*, 45–48

Call, H.P. and Mucke, I. (1997) History, overview and applications of mediated lignolytic systems, especially laccase-mediator-systems (Lignozym(R)-process). J. Biotechnol. *53*, 163–202

Camarero, S., Ibarra, D., Martinez, M.J. and Martinez, A.T. (2005) Lignin-derived compounds as efficient laccase mediators for decolorization of different types of recalcitrant dyes. Appl. Environ. Microbiol. *71*, 1775–1784

Chen, T, Barton SC., Binyamin G, Gao Z, Zhang Y, Kim HH and Heller A (2001) A miniature biofuel cell. J. Am. Chem. Soc. *123*, 8630–8631

Claus, H. (2003) Laccases and their occurrence in prokaryotes. Arch. Microbiol. *179*, 145–150

Claus, H. (2004) Laccases: structure, reactions, distribution. Micron *35*, 93–96

Collins, P.J., Kotterman, M.J.J., Field, J.A. and Dobson, A.D.W. (1996) Oxidation of anthracene and benzo[a]pyrene by laccases from *Trametes versicolor*. Appl. Environ. Microbiol *62*, 4563–4567

Davies, G.J.a.D. (2002) Laccase. Handbook of Metalloproteins. (ed A.Messerschmidt, *et al*.), pp. 1359–1368. John Wiley and Sons, LTD, New York

Ducros, V., Brzozowski, A.M., Wilson, K.S., Brown, S.H., Ostergaard, P., Schneider, P., Yaver, D.S., Pedersen, A.H. and Davies, G.J. (1998) Crystal structure of the type-2 Cu depleted laccase from *Coprinus cinereus* at 2.2 angstrom resolution. Nature Struct. Biol. *5*, 310–316

Duran, N., Rosa, M.A., D'Annibale, A. and Gianfreda, L. (2002) Applications of laccases and tyrosinases (phenoloxidases) immobilized on different supports: a review. Enzyme Microb. Technol. *31*, 907–931

Enguita, F.J., Martins, L.O., Henriques, A.O. and Carrondo, M.A. (2003) Crystal structure of a bacterial endospore coat component - A laccase with enhanced thermostability properties. J. Biol. Chem. *278*, 19416–19425

Felby, C., Hassingboe, J. and Lund, M. (2002) Pilot-scale production of fiberboards made by laccase oxidized wood fibers: board properties and evidence for cross-linking of lignin. Enzyme Microb. Technol. *31*, 736–741

Freire, R.S., Duran, N. and Kubota, L.T. (2002) Development of a laccase-based flow injection electrochemical biosensor for the determination of phenolic compounds and its application for monitoring remediation of Kraft E1 paper mill effluent. Anal. Chim. Acta *463*, 229–238

Freire, R.S., Ferreira, M.M.C., Duran, N. and Kubota, L.T.(2003) Dual amperometric biosensor device for analysis of binary mixtures of phenols by multivariate calibration using partial least squares. Anal. Chim. Acta *485*, 263–269

Gianfreda, L., Xu, F. and Bollag, J. (1999) Laccases: a useful group of oxidoreductive enzymes. Bioremed. J. *3*, 1–25

Hakulinen, N., Kiiskinen, L.L., Kruus, K., Saloheimo, M., Paananen, A., Koivula, A. and Rouvinen, J. (2002) Crystal structure of a laccase from *Melanocarpus albomyces* with an intact trinuclear copper site. Nature Struct. Biol. *9*, 601–605

Huttermann, A., Mai, C. and Kharazipour, A. (2001) Modification of lignin for the production of new compounded materials. Appl. Microbiol. Biotechnol. *55*, 387–394

Jauregui, J., Valderrama, B., Albores, A. and Vazquez-Duhalt, R. (2003) Microsomal transformation of organophosphorus pesticides by white rot fungi. Biodegradation *14*, 397–406

Johannes, C. and Majcherczyk, A. (2000) Natural mediators in the oxidation of polycyclic aromatic hydrocarbons by laccase mediator systems. Appl. Environ. Microbiol. *66*, 524–528

Johannes, C., Majcherczyk, A. and Huttermann, A. (1996) Degradation of anthracene by laccase of *Trametes versicolor* in the presence of different mediator compounds. Appl. Microbiol. Biotechnol. *46*, 313–317

Kiiskinen, L.L. and Saloheimo, M. (2004) Molecular cloning and expression in *Saccharomyces cerevisiae* of a laccase gene from the ascomycete *Melanocarpus albomyces*. Appl. Environ. Microbiol. *70*, pp. 137–144

Kumar, S.V.S., Phale, P.S., Durani, S. and Wangikar, P.P. (2003) Combined sequence and structure analysis of the fungal laccase family. Biotechnol. Bioeng. *83*, 386–394

Larsson, S., Cassland, P. and Jonsson, L.J. (2001) Development of a *Saccharomyces cerevisiae* strain with enhanced resistance to phenolic fermentation inhibitors in lignocellulose hydrolysates by heterologous expression of laccase. Appl Environ Microbiol *67*, 1163–1170

Luterek, J., Gianfreda, L., Wojtas-Wasilewska, M., Cho, N.S., Rogalski, J., Jaszek, H., Malarczyk, E., Staszczak, M. and Fink-Boots, M. (1998) Activity of free and immobilized extracellular Cerrena unicolor laccase in water miscible organic solvents. Holzforschung *52*, 589–595

Majcherczyk, A. and Johannes, C. (2000) Radical mediated indirect oxidation of a PEG-coupled polycyclic aromatic hydrocarbon (PAH) model compound by fungal laccase. Biochim. Biophys. Acta-Gen. Subj. *1474*, 157–162

Majcherczyk, A., Johannes, C. and Huttermann, A. (1998) Oxidation of polycyclic aromatic hydrocarbons (PAH) by laccase of *Trametes versicolor*. Enzyme Microb. Technol. *22*, 335–341

Mathiasen, T.E. Laccase and beer storage. [PCT Int.Appl.WO 9521240 A2]. 1995. Ref Type: Patent

Mayer, A.M. and Staples, R.C. (2002) Laccase: new functions for an old enzyme. Phytochemistry *60*, 551–565

Milligan, C. and Ghindilis, A. (2002) Laccase based sandwich scheme immunosensor employing mediatorless electrocatalysis. Electroanalysis *14*, 415–419

Minussi, R.C., Pastore, G.M. and Duran, N. (2002) Potential applications of laccase in the food industry. Trends Food Sci. Technol. *13*, 205–216

Nagai, M., Kawata, M., Watanabe, H., Ogawa, M., Saito, K., Takesawa, T., Kanda, K. and Sato, T. (2003) Important role of fungal intracellular laccase for melanin synthesis: purification and characterization of an intracellular laccase from *Lentinula edodes* fruit bodies. Microbiology-Sgm *149*, 2455–2462

Onuki, T., Nogucji, M. and Mitamura, J. Oxidative hair dye composition containing laccase. [WO 0037,030. Chem. Abstr. 133, 78994m]. 2000. Ref Type: Patent

Parkinson, N.M., Conyers, C.M., Keen, J.N., MacNicoll, A.D., Smith, I. and Weaver, R.J. (2003) cDNAs encoding large venom proteins from the parasitoid wasp *Pimpla hypochondriaca* identified by random sequence analysis. Comp. Biochem. Physiol. C-Toxicol. Pharmacol. *134*, 513–520

Pickard, M.A., Roman, R., Tinoco, R. and Vazquez-Duhalt, R. (1999) Polycyclic aromatic hydrocarbon metabolism by white rot fungi and oxidation by *Coriolopsis gallica* UAMH 8260 laccase. Appl. Environ. Microbiol. *65*, 3805–3809

Piontek, K., Antorini, M. and Choinowski, T. (2002) Crystal structure of a laccase from the fungus *Trametes versicolor* at 1.90-angstrom resolution containing a full complement of coppers. J. Biol. Chem. *277*, 37663–37669

Pruche, F., Saint, L.P. and Bernards, B. Hair dye compositions containing hydroxystilbene. [Eur.Pat. Appl. EP 1013,260. Chem. Abstr. 133,63587g.]. 2000. Ref Type: Patent

Rodakiewicz-Nowak, J., Kasture, S.M., Dudek, B. and Haber, J. (2000) Effect of various water-miscible solvents on enzymatic activity of fungal laccases. J. Mol. Catal. B-Enzymatic *11*, 1–11

Servili, M., De Stefano, G., Piacquadio, P. and Sciancalepore, V. (2000) A novel method for removing phenols from grape must. Amer. J. Enol. Viticult. *51*, 357–361

Shleev, S., Tkac, J., Christenson, A., Ruzgas, T., Yaropolov, A.I., Whittaker, J.W. and Gorton, L. (2005) Direct electron transfer between copper-containing proteins and electrodes. Biosen. Bioelectron. *20*, 2517–2554

Sigoillot, C., Camarero, S., Vidal, T., Record, E., Asther, M., Perez-Boada, M., Martinez, M.J., Sigoillot, J.C., Asther, M., Colom, J.F. and Martinez, A.T. (2005) Comparison of different fungal enzymes for bleaching high-quality paper pulps. J. Biotechnol. *115*, 333–343

Tetsch, L., Bend, J., Janssen, M. and Holker, U. (2005) Evidence for functional laccases in the acidophilic ascomycete *Hortaea acidophila* and isolation of laccase-specific gene fragments. FEMS Microbiol. Letters *245*, 161–168

Thurston, C.F. (1994) The structure and function of fungal laccases. Microbiology *140*, pp. 19–26

Torres, E., Bustos-Jaimes, I. and Le Borgne, S. (2003) Potential use of oxidative enzymes for the detoxification of organic pollutants. Appl. Catal. B-Environ. *46*, 1–15

Ullah, M.A., Bedford, C.T. and Evans, C.S. (2000) Reactions of pentachlorophenol with laccase from *Coriolus versicolor*. Appl. Microbiol. Biotechnol. *53*, 230–234

Valderrama, B., Oliver, P., Medrano-Soto, A. and Vazquez-Duhalt, R. (2003) Evolutionary and structural diversity of fungal laccases. Antonie Van Leeuwenhoek *84*, 289–299

Widsten, P., Tuominen, S., Qvintus-Leino, P. and Laine, J.E. (2004) The influence of high defibration temperature on the properties of medium-density fiberboard (MDF) made from laccase-treated softwood fibers. Wood Sci. Technol. *38*, 521–528

Xu, F., Berka, R.M., Wahleithner, J.A., Nelson, B.A., Shuster, J.R., Brown, S.H., Palmer, A.E. and Solomon, E.I. (1998) Site-directed mutations in fungal laccase: effect on redox potential, activity and pH profile. Biochem. J. *334*, 63–70

Yaropolov, A.I., Skorobogatko, O.V., Vartanov, S.S. and Varfolomeyev, S.D. (1994) laccase - properties, catalytic mechanism, and applicability. Appl. Biochem. Biotechnol. *49*, 257–280

CHAPTER 27

HIGH REDOX POTENTIAL PEROXIDASES

ÁNGEL T. MARTÍNEZ*
Centro de Investigaciones Biológicas, CSIC, Madrid, Spain
**atmartinez@cib.csic.es*

1. INTRODUCTION

High redox potential peroxidases secreted by some basidiomycete fungi are unique enzymes that play a central role in the degradation of lignin (Kirk and Cullen, 1998). Plant biomass (bryophytes excluded) is mainly made up of two polysaccharides, cellulose and hemicelluloses, and the complex aromatic polymer of lignin (Fengel and Wegener, 1984). The main role of lignin in the plant cell wall is the protection of polysaccharides against the action of hydrolytic enzymes. It also contributes to stem rigidity and water transport, two important characteristics that facilitated land colonization by vascular plants. Such characteristics derive from the recalcitrance and structural complexity of lignin which includes up to three types of subunit derived from three different *p*-hydroxycinnamyl alcohols (Higuchi, 1997; Boerjan *et al*, 2003). These phenylpropanoid units are linked together by a variety of ether and C-C bonds forming a three dimensional network that confers lignin its mechanical resistance as well as an extremely high resistance to degradation.

Biodegradation of the lignin polymer has been described as an "enzymatic combustion", where the aromatic units are oxidized by the hydrogen peroxide secreted by ligninolytic basidiomycetes in a reaction catalysed by high redox potential peroxidases, enzymes which are unique to this group of organisms (Kirk and Farrell, 1987). Due to the light colour of the wood decayed by ligninolytic basidiomycetes these organisms are also known as white-rot fungi (Martínez *et al*, 2005). By contrast, the so-called brown-rot basidiomycetes do not produce ligninolytic enzymes but instead transform wood into a lignin-enriched brown material. In addition to peroxidases, other fungal enzymes such as laccases and H_2O_2-producing oxidases are also involved in lignin biodegradation. Ligninolytic enzymes are of

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 477–488

considerable importance for the natural degradation of lignocellulosic materials in terrestrial ecosystems, enabling the recycling of the organic carbon fixed by photosynthesis, and also for most industrial processes that utilise lignocellulosic biomass, including paper pulp manufacturing and bioethanol production both of which require the removal of lignin to exploit cellulose.

2. LIGNINOLYTIC ENZYMES

Two ligninolytic peroxidases, lignin peroxidase (LiP) and manganese peroxidase (MnP), were discovered in 1983–84 in the white-rot fungus *Phanerochaete chrysosporium* (order Aphyllophorales) and described as "ligninases" because of their high redox potentials which enable the oxidation of dimeric lignin model compounds (Tien and Kirk, 1983; Glenn *et al*., 1983; Kuwahara *et al*., 1984). LiP is able to degrade non-phenolic lignin units (up to 90% of the polymer) or the simple compound 3,4-dimethoxybenzyl (veratryl) alcohol (Tien and Kirk, 1983), whereas MnP generates Mn^{3+} that, when chelated with oxalic or other organic acids, acts as a diffusible oxidizer of the phenolic lignin units (Glenn *et al*., 1986) or the non-phenolic units *via* lipid peroxidation reactions (Jensen *et al*., 1996). Versatile peroxidase (VP), a third type of ligninolytic peroxidase, was more recently described in *Pleurotus* and *Bjerkandera* species (from the orders Agaricales and Aphyllophorales, respectively) (Martínez *et al*., 1996; Mester and Field, 1998; Ruiz-Dueñas *et al*., 1999) and combines the catalytic properties of LiP, MnP and other peroxidases (from plants and micro-organisms) that oxidise phenolic compounds (Heinfling *et al*., 1998).

Laccases are known in plants, fungi, bacteria and insects, where they play a variety of roles including the synthesis of pigments, fruiting-body morphogenesis and detoxification (Thurston, 1994; Mayer and Staples, 2002) (see chapter by Alcalde in this book). Laccase production on solid media was considered to be a characteristic of white-rot fungi (Kääriäk, 1965), although some brown-rot fungi produce laccase in liquid cultures (Lee *et al*., 2004). However, these phenoloxidases have low redox potentials that only permit the oxidation of phenolic lignin units (often representing less than 10% of the polymer). Recently, biotechnological interest in laccases has been spurred by the discovery of their ability to oxidize high redox potential substrates in the presence of synthetic redox mediators, forming stable free radicals which act as diffusible oxidizers (Bourbonnais and Paice, 1990). The existence of natural mediators involved in lignin biodegradation has not been demonstrated despite various attempts (Li *et al*., 2001), although some lignin-derived phenols can act as laccase mediators (Camarero *et al*., 2005).

Other extracellular enzymes involved in lignin degradation are the H_2O_2-generating oxidases and mycelium-associated dehydrogenases. The former include aryl-alcohol oxidase (AAO), first described in *Pleurotus* and *Bjerkandera* species (Bourbonnais and Paice, 1988; Muheim *et al*., 1990; Guillén *et al*., 1992), and glyoxal oxidase found in *P. chrysosporium* and other basidiomycetes (Kersten, 1990). Fungal dehydrogenases, including aryl-alcohol dehydrogenase (AAD), are also involved in lignin degradation (Gutiérrez *et al*., 1994; Brock *et al*., 1995).

Ligninolytic peroxidases (LiP, MnP and VP) oxidize the lignin polymer in a one-electron reaction resulting in cation radical formation (phenoxy radicals from phenolic units) (Kersten *et al*., 1985). The aromatic cation radicals formed evolve through different non-enzymatic reactions resulting in breakdown of ether and C-C inter-unit linkages, demethoxylation and aromatic ring cleavage (Schoemaker, 1990; Martínez *et al*., 2005). The aromatic aldehydes released after oxidative degradation of the lignin-unit side-chains, or synthesized *de novo* by fungi (Gutiérrez *et al*., 1994), constitute the substrate for the generation of the H_2O_2 required for lignin degradation by peroxidases in *Pleurotus* species. In these and other white-rot fungi H_2O_2 is continuously formed in cyclic redox reactions involving both extracellular AAO and mycelial AAD (Guillén *et al*., 1994). Some oxidases and laccases also participate in the (direct or indirect) reduction of ferric to ferrous iron that reacts with H_2O_2 yielding the hydroxyl free radical, a very strong oxidizer involved in wood biodegradation (Evans *et al*., 1994; Guillén *et al*., 2000; Hammel *et al*., 2002). Ultimately, simple breakdown products from lignin degradation enter the fungal hyphae and are incorporated into the intracellular catabolic routes.

3. STRUCTURE AND FUNCTION OF LIGNINOLYTIC PEROXIDASES

Due to their uniqueness as industrial biocatalysts, the catalytic mechanisms of lignin-degrading peroxidases have been extensively investigated at the molecular level (Banci, 1997; Gold *et al*., 2000; Martínez, 2002). *Phanerochaete chrysosporium* LiP and MnP were the second and third peroxidases (after yeast cytochrome *c* peroxidase, CCP) whose crystal structures were solved, just ten years after their discovery (Poulos *et al*., 1993; Piontek *et al*., 1993; Sundaramoorthy *et al*., 1994); the crystal structure of the *Pleurotus eryngii* VP has been solved recently (Pérez-Boada *et al*., 2005; Piontek *et al*., 2006). As mentioned above, these peroxidases are able to catalyse the oxidation of the recalcitrant lignin units by H_2O_2. This is made possible by the formation of a high redox state oxo-ferryl intermediate after the two-electron reaction of the haem cofactor with H_2O_2. The activated enzyme oxidizes two substrate units and is reduced back to the resting state that reacts again with peroxide. This catalytic cycle, which includes the resting peroxidase and the so-called compound I (two-electron oxidized enzyme) and compound II (one-electron oxidized enzyme), is common to other peroxidases such as the well-known horseradish peroxidase (HRP) (Dunford, 1999). However, two aspects of their molecular structure provide the ligninolytic peroxidases their unique catalytic properties, namely: **i)** a haem environment conferring high redox potential to the oxo-ferryl complexes; and **ii)** specific binding sites (and mechanisms) for oxidation of their characteristic substrates including non-phenolic aromatic compounds in the case of LiP, manganous iron in the case of MnP, and both types of compounds in the case of VP. The hybrid properties of *Pleurotus eryngii* VP are illustrated in the basic catalytic cycle shown in Fig. 1A which combines the previously described cycles of LiP and MnP.

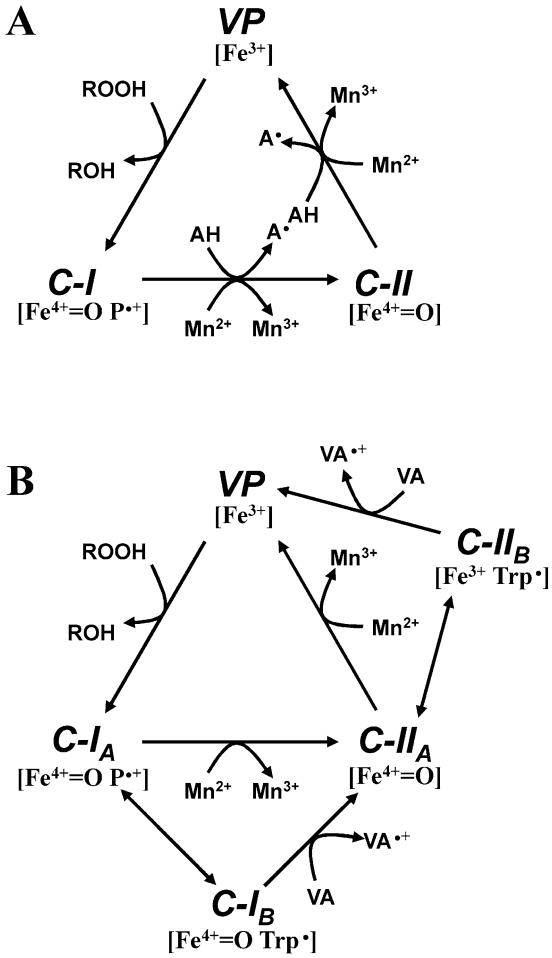


Figure 1. Catalytic cycle of the model ligninolytic peroxidase VP. **A**, Basic cycle as described by Ruiz-Dueñas *et al*. (1999), including two-electron oxidation of the resting peroxidase (VP) by hydroperoxide to yield compound-I (C-I), whose reduction in two one-electron reactions results in the intermediate compound-II (C-II) and then the resting form of the enzyme. VP can oxidize both aromatic substrates (AH) and Mn^{2+}. **B**, Extended cycle adapted from Pérez-Boada *et al*. (2005), including $C\text{-}I_B$ and $C\text{-}II_B$ involved in oxidation of veratryl alcohol (VA) and other high-redox potential aromatic substrates, where a tryptophan radical is present at position 164 corresponding to LiP W171 (low redox potential aromatic compounds are probably oxidized by both the A and B forms but for simplicity they are not included)

Similarities in the haem environment (located in the central region of the protein) between the three ligninolytic peroxidase types were evidenced by ^{1}H-NMR and confirmed by X-ray diffraction of peroxidase crystals (Banci, 1997; Banci *et al*., 2003). In NMR solution experiments it is possible to identify the signals of protons of the haem cofactor as well as those of several amino acid residues in the haem

pocket. This is possible due to the paramagnetic effects caused by the haem iron that displaces the signals of neighbouring protons outside the region where most protein protons overlap. These signals are better assigned in the spectra of peroxidase CN-adducts that yield hyperfine NMR signals. One of the main differences between the various peroxidases is the position of the proximal histidine side-chain Nε (acting as the fifth iron ligand) which in the ligninolytic peroxidases is displace away from the haem (compared to HRP or CCP) contributing to the electropositivity of the oxo-ferryl complex (Poulos, 1993; Banci, 1997; Martínez, 2002). The NMR results also explain the high redox potential of MnP, despite that this peroxidase is only able to steady-state oxidize Mn^{2+}, and suggest that this is due to the absence of a specific site in MnP for the oxidation of aromatic substrates, which is present in LiP and VP. Similarly, LiP lacks the Mn-binding site described below.

4. SUBSTRATE OXIDATION SITES IN LIGNINOLYTIC PEROXIDASES

Recent studies based on peroxidase molecular structures solved by crystal X-ray diffraction are contributing to the identification of the different substrate oxidation sites in ligninolytic peroxidases. The manganese and aromatic substrate binding-sites were first identified in MnP and LiP respectively (Doyle *et al*., 1998; Gold *et al*., 2000; Blodig *et al*., 2001), and then in VP after its crystal structure was solved. Studies on the molecular structure of VP confirmed that its novel catalytic properties are due to a hybrid molecular structure, as was suggested several years ago (Ruiz-Dueñas *et al*., 1999; Camarero *et al*., 1999).

Mn^{2+} oxidation by VP occurs at a binding site similar to that of MnP situated near one of the propionates of the haem group. The cation is bound by three acidic residues (one aspartate and two glutamates) enabling direct electron transfer to the cofactor (Fig. 2, left side). Initially it had been suggested that Mn^{2+} is oxidized by MnP at the edge of the haem in front of the classical access channel (δ-position following NMR nomenclature) (Harris *et al*., 1991). However, it has since been demonstrated that this cation uses a second access channel delimited by the three acidic residues mentioned above, situated in front of the most internal haem propionate (with respect to the main channel) (Sundaramoorthy *et al*., 1997). The functionality of this site in VP was confirmed by site-directed mutagenesis and the NMR spectra obtained during Mn^{2+} titration of the enzyme suggest that a unique Mn-binding site exists in the vicinity of the VP haem (Banci *et al*., 2003).

By contrast, it has been shown that veratryl alcohol and other lignin model substrates are oxidized by the ligninolytic peroxidases at the surface of the protein by a long-range electron transfer mechanism that initiates at an exposed tryptophan residue (Fig. 2, right side). The rationale for this mechanism is that most LiP and VP aromatic substrates, including the lignin polymer, cannot penetrate inside the protein to transfer electrons directly to the cofactor in the haem pocket. Thus, they are oxidized at the enzyme surface and the electrons travel to the haem by a

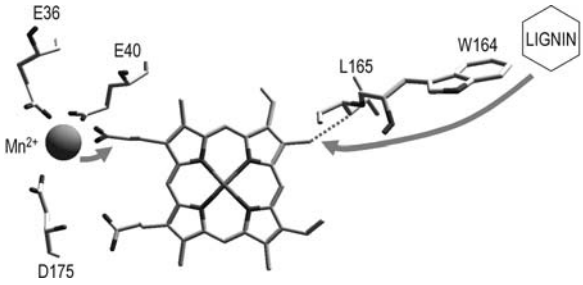


Figure 2. Detail of substrate oxidation sites in the model ligninolytic peroxidase VP including: **i)** Mn^{2+} oxidation site (left) where this cation is bound by the side-chains of E36, E40 and D175 and the internal propionate of haem that directly receives one electron from the substrate; and **ii)** high redox potential aromatic substrate (such as lignin and veratryl alcohol) oxidation site (right) where one electron travels *via* a long-range electron transfer pathway formed by an exposed tryptophan residue (W164, in the form of a tryptophanyl radical in the H_2O_2-activated VP) and a leucine residue connected to the haem (axial view from the proximal side of the haem group)

protein pathway. The tryptophan residue where the protein radical is expected to be produced was identified in LiP to be W171, which is involved in the oxidation of veratryl alcohol (Doyle *et al*., 1998) and a tetrameric lignin model compound (Mester *et al*., 2001). The formation of a protein radical was suggested by indirect evidence (Blodig *et al*., 1999). The homologous residue in VP (W164) has been identified, and two other putative long electron transfer pathways proposed for LiP have been discounted by site-directed mutagenesis studies. Simultaneously, the existence of a tryptophanyl protein radical was directly detected for the first time in a ligninolytic peroxidase using low-temperature EPR of H_2O_2-activated VP (Pérez-Boada *et al*., 2005; Pogni *et al*., 2005). Moreover, using a combination of multi-frequency EPR and DFT (density-functional theory) calculations it has been possible to identify the VP protein radical as a neutral tryptophanyl radical and confirm that it corresponds to W164 as suggested by site-directed mutagenesis (Pogni *et al*., 2006). On the other hand, it has been found that W171 in wild LiP (isoenzymes H2 and H8), as well as in *Escherichia coli*-expressed LiP treated with several H_2O_2 equivalents, is C_β-hydroxylated (Blodig *et al*., 1998; Choinowski *et al*., 1999, 2001). This finding was related to the formation of the above-mentioned protein radical, however, no trace of W164 hydroxylation was found in the atomic resolution structures of recombinant VP treated with H_2O_2 or in wild VP (Pérez-Boada *et al*., 2005). This result demonstrates that hydroxylation is not an immediate consequence of the formation of the tryptophan radical in ligninolytic peroxidases (the evolution of this reactive radical in the absence of reducing substrate probably depends on its protein environment that is different in VP and LiP).

Finally, several pieces of evidence suggest that similarly to HRP (Henriksen *et al*., 1999) and the low redox potential fungal peroxidase from litter-decomposing *Coprinus* species (Tsukamoto *et al*., 1999), some low redox potential substrates –

including substituted phenols and dyes like 2, 2′-azinobis-(3-ethylbenzothiazoline-6-sulfonic acid) – can be oxidized at the main haem access channel of VP and in some cases of LiP (Doyle *et al*., 1998).

In light of the above information the basic cycle of VP (Fig. 1A), as a model ligninolytic peroxidase, can be now completed in the part corresponding to the oxidation of veratryl alcohol and other high redox potential substrates (Fig. 1B). This modification, which also agrees with the mechanisms proposed in LiP, implies the existence of two additional peroxidase forms, compounds I_B and II_B, where one oxidation equivalent would be in the form of a tryptophan radical. These two forms would be in equilibrium with the classical intermediate compounds (now I_A and II_A), where the oxidation equivalent is in the form of a porphyrin cation radical and a Fe^{4+}-oxo, respectively. The percentage of forms A and B in the two equilibrium reactions could be variable, and their involvement in VP catalysis will depend on the substrate. Thus the "classical" compounds I_A and II_A will be predominant during oxidation of Mn^{2+}, whereas a certain percentage of the protein radical forms will be present during oxidation of veratryl alcohol. EPR results have shown that VP compound I_B (including W164 radical) is present in the compound I equilibrium (representing around 25% under the experimental conditions used) (Pérez-Boada *et al*., 2005; Pogni *et al*., 2006). The existence of compound II_B, which in CCP represents 10% of total compound II at pH 5 (Ho *et al*., 1983), is still to be directly demonstrated in VP, though its presence is required for catalysis as explained above.

5. INDUSTRIAL BIOCATALYSTS

The use of enzymes as industrial biocatalysts to remove lignin-derived compounds responsible for the dark colour of paper pulps was introduced several years ago to substitute chlorine-containing bleaching reagents that cause negative environmental impact (Viikari *et al*., 1994; Bajpai, 2004). The first enzymes used in the bleach plant were hemicellulose-degrading xylanases (see Chapter 5 of section A) since the positive effects of xylanases in pulp bleaching were first shown (Kantelinen *et al*., 1990) at a time when barely any information was available on ligninolytic enzymes. However, xylanases have limited efficiency in bleaching as they only contribute to remove natural xylan-lignin complexes and lignin fragments that co-precipitate with xylan after cooking. Thus, xylanases are mainly used in so-called elementary chlorine free (ECF) bleaching in combination with chlorine dioxide. In order to develop new enzyme-based totally chlorine free (TCF) sequences the use of lignin-degrading enzymes is required.

In addition to xylanases, ligninolytic enzymes have been being studied in recent years for pulp bleaching application and include the use of both peroxidases and laccases (Paice *et al*., 1995; Bajpai, 2004; Sigoillot *et al*., 2005). Laccases have been known for decades and are already commercially produced in high yields using heterologous expression systems. By contrast, ligninolytic peroxidases are not currently commercially available despite considerable efforts devoted to improve

their expression in fungal hosts (Conesa *et al*., 2000, 2002; Gu *et al*., 2003; Lú-Chau *et al*., 2004). As already mentioned, laccases have a low redox potential and can degrade lignin only in the presence of synthetic redox mediators, the so-called laccase-mediator system (see Chapter by Alcalde in this volume). However, it is important to note that it is more difficult to reduce the costs associated with obtaining the chemically synthesised mediators than those of the enzymes obtained by genetic engineering. This together with the potential risks of mediator release to the environment, are currently the main obstacles to the use of laccase-mediator systems at pulp mill scale.

By contrast, ligninolytic peroxidases do not require mediators to degrade high redox-potential compounds which represents a significant advantage compared to laccases. Indeed, they should be the enzymes of choice for removing lignin or transforming high-redox potential aromatic compounds in different industrial applications. Hence the rapid acquisition of knowledge on the structure and function of these enzymes over the last few years. The information currently available has been used to modulate their catalytic (substrate specificity) and operational properties (including enzyme stability) using site-directed mutagenesis as illustrated in various studies (Yeung *et al*., 1997; Wilcox *et al*., 1998; Timofeevski *et al*., 1999; Reading and Aust, 2000; Celik *et al*., 2001; Mester and Tien, 2001; Feng *et al*., 2003). In those cases where the structural basis of the property to be improved is unknown, or too difficult to be predicted, directed evolution is the approach of choice (Cherry *et al*., 1999; MiyazakiImamura *et al*., 2003). Moreover, gene fusion permits the introduction of cellulose binding domains to improve peroxidase efficiency (Levy *et al*., 2003) or the generation of functional multi-enzymatic systems. It is to be expected that genetic engineering techniques will also contribute to overcoming the main drawback for the industrial utilization of ligninolytic peroxidases, *i.e.* their low yields in heterologous expression. In addition to increasing the expression levels of the target genes (*e.g.* by changing promoters or signal peptides, gene co-expression or fusion, *etc*.) an alternative approach that is currently being explored is to engineer an already commercial (*i.e.* high-yield expressed) peroxidase by introducing the ability to oxidize high-redox potential compounds by protein engineering. Using a combination of site-directed mutagenesis, directed evolution and other techniques such as chemical stabilization, tailor-made biocatalysts will be available in the future enabling industrial exploitation of the unique characteristics of the high redox peroxidases produced by lignin-degrading fungi.

ACKNOWLEDGEMENTS

Recent lignin biodegradation studies at the CIB were supported by ENCE (Spain), Spanish projects BIO2002-1166 and BIO2005-03569, and the EU projects QLK3-99-590 and NMP2-CT-2006-026456. The author thanks K. Piontek (ETH, Zurich) for the unpublished crystal structure of VP, R. Pogni (University of Sienna) for EPR studies, and F.J. Ruiz-Dueñas (CIB, Madrid) for valuable comments.

REFERENCES

Bajpai, P. (2004). Biological bleaching of chemical pulps. Crit. Rev. Biotechnol. *24*, 1–58.

Banci, L. (1997). Structural properties of peroxidases. J. Biotechnol. *53*, 253–263.

Banci, L., Camarero, S., Martínez, A.T., Martínez, M.J., Pérez-Boada, M., Pierattelli, R. and Ruiz-Dueñas, F.J. (2003). NMR study of Mn(II) binding by the new versatile peroxidase from the white-rot fungus *Pleurotus eryngii*. J. Biol. Inorg. Chem. *8*, 751–760.

Blodig, W., Doyle, W.A., Smith, A.T., Winterhalter, K., Choinowski, T.H. and Piontek, K. (1998). Autocatalytic formation of hydroxy group at Cβ of Trp171 in lignin peroxidase. Biochemistry *37*, 8832–8838.

Blodig, W., Smith, A.T., Doyle, W.A. and Piontek, K. (2001). Crystal structures of pristine and oxidatively processed lignin peroxidase expressed in *Escherichia coli* and of the W171F variant that eliminates the redox active tryptophan 171. Implications for the reaction mechanism. J. Mol. Biol. *305*, 851–861.

Blodig, W., Smith, A.T., Winterhalter, K. and Piontek, K. (1999). Evidence from spin-trapping for a transient radical on tryptophan residue 171 of lignin peroxidase. Arch. Biochem. Biophys. *370*, 86–92.

Boerjan, W., Ralph, J. and Baucher, M. (2003). Lignin biosynthesis. Annu. Rev. Plant Biol. *54*, 519–546.

Bourbonnais, R. and Paice, M.G. (1988). Veratryl alcohol oxidases from the lignin degrading basidiomycete *Pleurotus sajor-caju*. Biochem. J. *255*, 445–450.

Bourbonnais, R. and Paice, M.G. (1990). Oxidation of non-phenolic substrates. An expanded role for laccase in lignin biodegradation. FEBS Lett. *267*, 99–102.

Brock, B.J., Rieble, S. and Gold, M.H. (1995). Purification and characterization of a 1,4-benzoquinone reductase from the basidiomycete *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. *61*, 3076–3081.

Camarero, S., Ibarra, D., Martínez, M.J. and Martínez, A.T. (2005). Lignin-derived compounds as efficient laccase mediators for decolorization of different types of recalcitrant dyes. Appl. Environ. Microbiol. *71*, 1775–1784.

Camarero, S., Sarkar, S., Ruiz-Dueñas, F.J., Martínez, M.J. and Martínez, A.T. (1999). Description of a versatile peroxidase involved in natural degradation of lignin that has both Mn-peroxidase and lignin-peroxidase substrate binding sites. J. Biol. Chem. *274*, 10324–10330.

Celik, A., Cullis, P.M., Sutcliffe, M.J., Sangar, R. and Raven, E.L. (2001). Engineering the active site of ascorbate peroxidase. Eur. J. Biochem. *268*, 78–85.

Cherry, J.R., Lamsa, M.H., Schneider, P., Vind, J., Svendsen, A., Jones, A. and Pedersen, A.H. (1999). Directed evolution of a fungal peroxidase. Nature Biotechnol. *17*, 379–384.

Choinowski, T., Blodig, W., Winterhalter, K. and Piontek, K. (1999). The crystal structure of lignin peroxidase at 1.70 Å resolution reveals a hydroxyl group on the C_β of tryptophan *171*, A novel radical site formed during redox cycle. J. Mol. Biol. *286*, 809–827.

Conesa, A., Jeenes, D., Archer, D.B., vandenHondel, C.A.M.J. and Punt, P.J. (2002). Calnexin overexpression increases manganese peroxidase production in *Aspergillus niger*. Appl. Environ. Microbiol. *68*, 846–851.

Conesa, A., van den Hondel, C.A.M.J.J. and Punt, P.J. (2000). Studies on the production of fungal peroxidases in *Aspergillus niger*. Appl. Environ. Microbiol. *66*, 3016–3023.

Doyle, W.A., Blodig, W., Veitch, N.C., Piontek, K. and Smith, A.T. (1998). Two substrate interaction sites in lignin peroxidase revealed by site-directed mutagenesis. Biochemistry *37*, 15097–15105.

Dunford, H.B. (1999). Heme peroxidases. Wiley-VCH, New York.

Evans, C.S., Dutton, M.V., Guillén, F. and Veness, R.G. (1994). Enzymes and small molecular mass agents involved with lignocellulose degradation. FEMS Microbiol. Rev. *13*, 235–240.

Feng, M.L., Tachikawa, H., Wang, X.T., Pfister, T.D., Gengenbach, A.J. and Lu, Y. (2003). Resonance Raman spectroscopy of cytochrome c peroxidase variants that mimic manganese peroxidase. J. Biol. Inorg. Chem. *8*, 699–706.

Fengel, D. and Wegener, G. (1984). Wood: Chemistry, ultrastructure, reactions. De Gruyter, Berlin.

Glenn, J.K., Akileswaran, L. and Gold, M.H. (1986). Mn(II) oxidation is the principal function of the extracellular Mn-peroxidase from *Phanerochaete chrysosporium*. Arch. Biochem. Biophys. *251*, 688–696.

Glenn, J.K., Morgan, M.A., Mayfield, M.B., Kuwahara, M. and Gold, M.H. (1983). An extracellular H_2O_2-requiring enzyme preparation involved in lignin biodegradation by the white rot basidiomycete *Phanerochaete chrysosporium*. Biochem. Biophys. Res. Commun. *114*, 1077–1083.

Gold, M.H., Youngs, H.L. and Gelpke, M.D. (2000). Manganese peroxidase. Met. Ions Biol. Syst. *37*, 559–586.

Gu, L., Lajoie, C. and Kelly, C. (2003). Expression of a *Phanerochaete chrysosporium* manganese peroxidase gene in the yeast *Pichia pastoris*. Biotechnol. Progr. *19*, 1403–1409.

Guillén, F., Gómez-Toribio, V., Martínez, M.J. and Martínez, A.T. (2000). Production of hydroxyl radical by the synergistic action of fungal laccase and aryl alcohol oxidase. Arch. Biochem. Biophys. *383*, 142–147.

Guillén, F., Martínez, A.T. and Martínez, M.J. (1992). Substrate specificity and properties of the aryl-alcohol oxidase from the ligninolytic fungus *Pleurotus eryngii*. Eur. J. Biochem. *209*, 603–611.

Guillén, F., Martínez, A.T., Martínez, M.J. and Evans, C.S. (1994). Hydrogen peroxide-producing system of *Pleurotus eryngii* involving the extracellular enzyme aryl-alcohol oxidase. Appl. Microbiol. Biotechnol. *41*, 465–470.

Gutiérrez, A., Caramelo, L., Prieto, A., Martínez, M.J. and Martínez, A.T. (1994). Anisaldehyde production and aryl-alcohol oxidase and dehydrogenase activities in ligninolytic fungi from the genus *Pleurotus*. Appl. Environ. Microbiol. *60*, 1783–1788.

Hammel, K.E., Kapich, A.N., Jensen, K.A. Jr. and Ryan, Z.C. (2002). Reactive oxygen species as agents of wood decay by fungi. Enzyme Microb. Technol *30*, 445–453.

Harris, R.Z., Wariishi, H., Gold, M.H. and Ortiz de Montellano, P.R. (1991). The catalytic site of manganese peroxidase. Regiospecific addition of sodium azide and alkylhydrazines to the heme group. J. Biol. Chem. *266*, 8751–8758.

Heinfling, A., Ruiz-Dueñas, F.J., Martínez, M.J., Bergbauer, M., Szewzyk, U. and Martínez, A.T. (1998). A study on reducing substrates of manganese-oxidizing peroxidases from *Pleurotus eryngii* and *Bjerkandera adusta*. FEBS Lett. *428*, 141–146.

Henriksen, A., Smith, A.T. and Gajhede, M. (1999). The structures of the horseradish peroxidase C-ferulic acid complex and the ternary complex with cyanide suggest how peroxidases oxidize small phenolic substrates. J. Biol. Chem. *274*, 35005–35011.

Higuchi, T. (1997). Biochemistry and Molecular Biology of Wood. Springer Verlag, London.

Ho, P.S., Hoffman, B.M., Kang, C.H. and Margoliash, E. (1983). Control of the transfer of oxidizing equivalents between heme iron and free radical site in yeast cytochrome *c* peroxidase. J. Biol. Chem. *258*, 4356–4363.

Jensen, K.A. Jr., Bao, W., Kawai, S., Srebotnik, E. and Hammel, K.E. (1996). Manganese-dependent cleavage of non-phenolic lignin structures by *Ceriporiopsis subvermispora* in the absence of lignin peroxidase. Appl. Environ. Microbiol. *62*, 3679–3686.

Käärik, A. (1965). The identification of the mycelia of wood-decay fungi by their oxidation reactions with phenolic compounds. Stud. For. Suec. *31*, 3–81.

Kantelinen, A., Viikari, L., Ranua, M., Sundquist, J. and Linko, M. (1990). Enzymatic bleaching. Kemia-Kemi *17*, 962.

Kersten, P.J. (1990). Glyoxal oxidase of *Phanerochaete chrysosporium*: Its characterization and activation by lignin peroxidase. Proc. Natl. Acad. Sci. USA *87*, 2936–2940.

Kersten, P.J., Tien, M., Kalyanaraman, B. and Kirk, T.K. (1985). The ligninase of *Phanerochaete chrysosporium* generates cation radicals from methoxybenzenes. J. Biol. Chem. *260*, 2609–2612.

Kirk, T.K. and Cullen, D. (1998). Enzymology and molecular genetics of wood degradation by white-rot fungi, p. 273–308. *In* R.A. Young and M. Akhtar (eds.), Environmentally friendly technologies for the pulp and paper industry. TAPPI Press, Atlanta.

Kirk, T.K. and Farrell, R.L. (1987). Enzymatic "combustion": The microbial degradation of lignin. Annu. Rev. Microbiol. *41*, 465–505.

Kuwahara, M., Glenn, J.K., Morgan, M.A. and Gold, M.H. (1984). Separation and characterization of two extracellular H_2O_2-dependent oxidases from ligninolytic cultures of *Phanerochaete chrysosporium*. FEBS Lett. *169*, 247–250.

Lee, K.H., Wi, S.G., Singh, A.P. and Kim, Y.S. (2004). Micromorphological characteristics of decayed wood and laccase produced by the brown-rot fungus *Coniophora puteana*. J. Wood Sci. *50*, 281–284.

Levy, I., Ward, G., Hadar, Y., Shoseyov, O. and Dosoretz, C.G. (2003). Oxidation of 4-bromophenol by the recombinant fused protein cellulose-binding domain-horseradish peroxidase immobilized on cellulose. Biotechnol. Bioeng. *82*, 223–231.

Li, K.C., Horanyi, P.S., Collins, R., Phillips, R.S. and Eriksson, K.E.L. (2001). Investigation of the role of 3-hydroxyanthranilic acid in the degradation of lignin by white-rot fungus *Pycnoporus cinnabarinus*. Enzyme Microb. Technol. *28*, 301–307.

Lú-Chau, T.A., Ruiz-Dueñas, F.J., Camarero, S., Feijoo, G., Martínez, M.J., Lema, J.M. and Martínez, A.T. (2004). Effect of pH on the stability of *Pleurotus eryngii* versatile peroxidase during heterologous production in *Emericella nidulans*. Bioprocess. Biosyst. Eng *26*, 287–293.

Martínez, A.T. (2002). Molecular biology and structure-function of lignin-degrading heme peroxidases. Enzyme Microb. Technol. *30*, 425–444.

Martínez, A.T., Speranza, M., Ruiz-Dueñas, F.J., Ferreira, P., Camarero, S., Guillén, F., Martínez, M.J., Gutiérrez, A. and del Río, J.C. (2005). Biodegradation of lignocellulosics: Microbiological, chemical and enzymatic aspects of fungal attack to lignin. Intern. Microbiol. *8*, 195–204.

Martínez, M.J., Ruiz-Dueñas, F.J., Guillén, F. and Martínez, A.T. (1996). Purification and catalytic properties of two manganese-peroxidase isoenzymes from *Pleurotus eryngii*. Eur. J. Biochem. *237*, 424–432.

Mayer, A.M. and Staples, R.C. (2002). Laccase: new functions for an old enzyme. Phytochemistry *60*, 551–565.

Mester, T., Ambert-Balay, K., Ciofi-Baffoni, S., Banci, L., Jones, A.D. and Tien, M. (2001). Oxidation of a tetrameric nonphenolic lignin model compound by lignin peroxidase. J. Biol. Chem. *276*, 22985–22990.

Mester, T. and Field, J.A. (1998). Characterization of a novel manganese peroxidase-lignin peroxidase hybrid isozyme produced by *Bjerkandera* species strain BOS55 in the absence of manganese. J. Biol. Chem. *273*, 15412–15417.

Mester, T. and Tien, M. (2001). Engineering of a manganese-binding site in lignin peroxidase isozyme H8 from *Phanerochaete chrysosporium*. Biochem. Biophys. Res. Commun. *284*, 723–728.

Miyazaki-Imamura, C., Oohira, K., Kitagawa, R., Nakano, H., Yamane, T. and Takahashi, H. (2003). Improvement of H_2O_2 stability of manganese peroxidase by combinatorial mutagenesis and high-throughput screening using *in vitro* expression with protein disulfide isomerase. Protein Eng. *16*, 423–428.

Muheim, A., Waldner, R., Leisola, M.S.A. and Fiechter, A. (1990). An extracellular aryl-alcohol oxidase from the white-rot fungus *Bjerkandera adusta*. Enzyme Microb. Technol. *12*, 204–209.

Paice, M.G., Bourbonnais, R., Reid, I.D., Archibald, F.S. and Jurasek, L. (1995). Oxidative bleaching enzymes: A review. J. Pulp Paper Sci. *21*, J280-J284.

Pérez-Boada, M., Ruiz-Dueñas, F.J., Pogni, R., Basosi, R., Choinowski, T., Martínez, M.J., Piontek, K. and Martínez, A.T. (2005). Versatile peroxidase oxidation of high redox potential aromatic compounds: Site-directed mutagenesis, spectroscopic and crystallographic investigations of three long-range electron transfer pathways. J. Mol. Biol. *345*, 385–402.

Piontek, K., Choinowski, T., Ruiz-Dueñas, F.J., Martínez, M.J. and Martínez, A.T. (2006). High-resolution crystal structure of versatile peroxidase oxidizing both Mn(II) and aromatic substrates. J. Mol. Biol. (in preparation).

Piontek, K., Glumoff, T. and Winterhalter, K. (1993). Low pH crystal structure of glycosylated lignin peroxidase from *Phanerochaete chrysosporium* at 2.5 Å resolution. FEBS Lett. *315*, 119–124.

Pogni, R., Baratto, M.C., Giansanti, S., Teutloff, C., Verdin, J., Valderrama, B., Lendzian, F., Lubitz, W., Vazquez-Duhalt, R. and Basosi, R. (2005). Tryptophan-based radical in the catalytic mechanism of versatile peroxidase from *Bjerkandera adusta*. Biochemistry *44*, 4267–4274.

Pogni, R., Baratto, M.C., Teutloff, C., Giansanti, S., Ruiz-Dueñas, F.J., Choinowski, T., Piontek, K., Martínez, A.T., Lendzian, F. and Basosi, R. (2006). A tryptophan neutral radical in the oxidized state of versatile peroxidase from *Pleurotus eryngii*: a combined multi-frequency EPR and DFT study. J. Biol. Chem. Jan 27 (Epub ahead of print).

Poulos, T.L. (1993). Peroxidases. Curr. Opin. Biotechnol. *4*, 484–489.

Poulos, T.L., Edwards, S.L., Wariishi, H. and Gold, M.H. (1993). Crystallographic refinement of lignin peroxidase at 2 Å. J. Biol. Chem. *268*, 4429–4440.

Reading, N.S. and Aust, S.D. (2000). Engineering a disulfide bond in recombinant manganese peroxidase results in increased thermostability. Biotechnol. Progr. *16*, 326–333.

Ruiz-Dueñas, F.J., Martínez, M.J. and Martínez, A.T. (1999). Molecular characterization of a novel peroxidase isolated from the ligninolytic fungus *Pleurotus eryngii*. Mol. Microbiol. *31*, 223–236.

Schoemaker, H.E. (1990). On the chemistry of lignin degradation. Recueil des Travaux Chimiques des Pays-Bas *109*, 255–272.

Sigoillot, C., Camarero, S., Vidal, T., Record, E., Asther, M., Pérez-Boada, M., Martínez, M.J., Sigoillot, J.-C., Asther, M., Colom, J. and Martínez, A.T. (2005). Comparison of different fungal enzymes for bleaching high-quality paper pulps. J. Biotechnol. *115*, 333–343.

Sundaramoorthy, M., Kishi, K., Gold, M.H. and Poulos, T.L. (1994). The crystal structure of manganese peroxidase from *Phanerochaete chrysosporium* at 2.06-Å resolution. J. Biol. Chem. *269*, 32759–32767.

Sundaramoorthy, M., Kishi, K., Gold, M.H. and Poulos, T.L. (1997). Crystal structures of substrate binding site mutants of manganese peroxidase. J. Biol. Chem. *272*, 17574–17580.

Thurston, C.F. (1994). The structure and function of fungal laccases. Microbiology *140*, 19–26.

Tien, M. and Kirk, T.K. (1983). Lignin-degrading enzyme from the hymenomycete *Phanerochaete chrysosporium* Burds. Science *221*, 661–663.

Timofeevski, S.L., Nie, G., Reading, N.S. and Aust, S.D. (1999). Engineering a functional hybrid of manganese peroxidase and lignin peroxidase. Plant Peroxidase Newsletter *13*, 99–111.

Tsukamoto, K., Itakura, H., Sato, K., Fukuyama, K., Miura, S., Takahashi, S., Ikezawa, H. and Hosoya, T. (1999). Binding of salicylhydroxamic acid and several aromatic donor molecules to *Arthromyces ramosus* peroxidase, investigated by X-ray crystallography, optical difference spectroscopy, NMR relaxation, molecular dynamics and kinetics. Biochemistry *38*, 12558–12568.

Viikari, L., Kantelinen, A., Sundquist, J. and Linko, M. (1994). Xylanases in Bleaching – From an Idea to the Industry. FEMS Microbiol. Rev. *13*, 335–350.

Wilcox, S.K., Putnam, C.D., Sastry, M., Blankenship, J., Chazin, W.J., McRee, D.E. and Goodin, D.B. (1998). Rational design of a functional metalloenzyme: Introduction of a site for manganese binding and oxidation into a heme peroxidase. Biochemistry *37*, 16853–16862.

Yeung, B.K.S., Wang, X.T., Sigman, J.A., Petillo, P.A. and Lu, Y. (1997). Construction and characterization of a manganese-binding site in cytochrome c peroxidase: Towards a novel manganese peroxidase. Chem. Biol. *4*, 215–221.

CHAPTER 28

AMINO ACID DEHYDROGENASES

STEPHEN Y.K. SEAH*

Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada
**sseah@uoguelph.ca*

1. INTRODUCTION

Amino acid dehydrogenases (EC 1.4.1) are enzymes that catalyze the reversible oxidative deamination of amino acids to form the corresponding keto-acids, using NAD^+ or $NADP^+$ as cofactors. This chapter will focus on the homologous glutamate, leucine, valine and phenylalanine dehydrogenases that catalyze the oxidative deamination of the α-amino of the substrate, represented by the general equation:

$$\mathrm{HOOC-\underset{\displaystyle H}{\overset{\displaystyle R}{\overset{|}{\underset{|}{C}}}}-NH_2 + NAD(P)^+ + H_2O \rightleftarrows HOOC-\overset{\displaystyle R}{\overset{|}{C}}=O + NH_4^+ + NAD(P)H}$$

Glutamate dehydrogenase (GluDH) is found in all three domains of life, being important for ammonia assimilation and for bridging carbon and nitrogen metabolism (Hudson and Daniel, 1993). Valine dehydrogenase (ValDH) was first reported in pea shoots and was subsequently found in members of the genus *Streptomyces*. ValDH is involved in the biosynthesis of *n*-butyrate which is a precursor of macrolide and polyether antibiotics in these bacteria (Tang *et al*, 1994; Leiser *et al*, 1996). ValDH has also been isolated from an Antarctic psychrophile *Cytophaga* sp. KUC-1 (Oikawa *et al.*, 2001). Both Leucine (LeuDH) and phenylalanine dehydrogenases (PheDH) are usually restricted to Gram positive aerobic bacteria. Their physiological roles are not clear, although it has been suggested that they are involved in catabolism of their respective preferred amino acid substrates to provide cellular energy and nitrogen (Asano *et al*, 1987a; Massey *et al*, 1976). GluDH from eubacteria and fungi are usually NAD^+ or $NADP^+$ specific. The vertebrate

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 489–504

and some archaeabacteria GluDH can use either coenzyme with equal efficacy. LeuDH, ValDH and PheDH from Bacteria are all NAD^+ specific.

Due to the stereospecific reaction catalyzed by the amino acid dehydrogenases, they are useful for the synthesis of amino acids for dietary or pharmaceutical purposes. The enzymes can also be used for quantitative determination of amino acids and ammonia in clinical applications. Members of this family of enzymes share sequence and structural similarity but may differ in terms of substrate specificity, stability, salt tolerance *etc*. GluDH, for example, can be found in organisms occupying different ecological niches and the enzymes are therefore "adapted" to different environmental conditions. Comparative structure-function studies over the years, beginning with the elucidation of the structure of GluDH from *Clostridium symbiosum* in 1985 (Rice *et al*, 1985), have provided a comprehensive understanding of the molecular basis of substrate recognition and thermostability in this family of enzymes. These are important parameters to consider in the biotechnological application of enzymes and deciphering the molecular determinants of such properties offer the potential to engineer desired characteristics in these biocatalysts.

2. GENERAL STRUCTURE AND REACTION MECHANISM

GluDH from the gram-positive glutamate fermenting micro-organism, *Clostridium symbiosum*, is the first member of the amino acid dehydrogenases with a structure solved by X-ray crystallography (Rice *et al*, 1985). Structures of other GluDHs (Bhuiya *et al*, 2005; Britton *et al*, 1999; Knapp *et al*, 1997; Peterson and Smith, 1999; Smith *et al*, 2002; Werner *et al*, 2005; Yip *et al*, 1995), a LeuDH (Baker *et al*, 1995) and a PheDH (Vanhooke *et al*, 1999) have since been determined and they all adopt a generally conserved structural fold. Minimally, each enzyme subunit contains two domains, I and II. Domain I is formed largely from residues in the N-terminal portion of the polypeptide chain and is involved in intersubunit interaction in homo-oligomeric structures. Domain II possesses an α/β fold reminiscent of a Rossmann fold, except that one of the β-strand's direction is reversed. This domain is responsible for binding of the coenzyme. The amino acid substrate is bound at the deep cleft between the two domains (Stillman *et al*, 1993). Bovine and human GluDHs have an extra "antenna domain" extending from Domain II (Peterson and Smith, 1999; Smith *et al*, 2002). This domain, which forms extensive contacts with adjacent subunits in the hexameric enzymes, are proposed to mediate allosteric regulation and negative cooperativity observed in these mammalian GluDHs (Fig. 1).

GluDHs are usually either hexameric or tetrameric (Britton *et al*, 1992). The other dehydrogenases have been reported to exhibit a variety of quaternary structures, from monomers to dodecamers (Brunhuber and Blanchard, 1994). Direct crystallographic structural analysis however, shows that *Bacillus sphaericus* LeuDH (Turnbull *et al*, 1997) and *Rhodococcus* sp. M4 PheDH (Vanhooke *et al*, 1999) are octameric and dimeric, respectively.

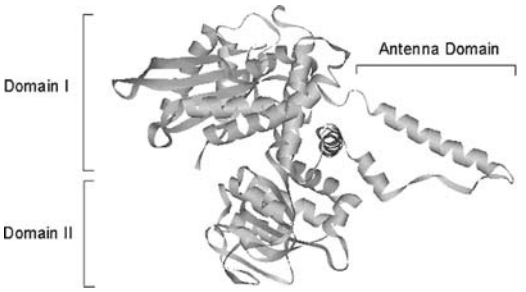


Figure 1. Structure of a single subunit of Human GluDH (Smith *et al*., 2002) showing its three characteristic enzyme domains

The general reaction mechanism of the enzymes in the reductive deamination direction, first proposed by Rife and Celand (1980) for GluDH, involves the deprotonation of the α-amino group of the substrate followed by the addition of an enzyme bound water molecule to form a carbinolamine intermediate. The removal of the amine group of the substrate as ammonia followed by proton rearrangement led to the formation of the keto acid. The residues aspartate and lysine (D165 and K125 in the Clostridial GluDH) within the active site is proposed to be the catalytic residues. This is supported by extensive site-specific mutagenesis studies on GluDHs and other amino acid dehydrogenases in the family (Dean *et al*., 1994; Hayden *et al*., 1999; Kataoka *et al*., 1994; Sekimoto *et al*., 1993; Hyun *et al*., 2000). Whether deprotonation of the amino group of the substrate is mediated by D165 or K125, however, remains to be resolved (Stillman *et al*., 1993; Brunhuber *et al*., 2000).

3. FUNCTION

3.1. Temperature Stability

Crystal structures have been solved for the GluDH from mesophiles (*Clostridium symbiosum, Escherichia coli*), thermophiles *(Thermotoga maritima*) and hyperthermophiles (*Pyrococcus furiosus* and *Thermococcus litoralis*). The general structures of these hexameric enzymes are similar although, as observed in other enzymes, there is an increase in the number of ion pairs networks correlating with increased thermostability (Yip *et al*., 1995). Increased ionic interactions by themselves do not necessarily lead to increased thermostability due to entropic and enthalpic penalties resulting from reduced mobility of the interacting residues and the loss of the more favourable hydration energies of charged residues. The enthalpic cost can be offset by the reduced hydration energy at higher temperatures (Elcock, 1998), while formation of clusters of ion pair networks appears to be the strategy adopted by the thermophilic GluDHs to overcome the unfavourable entropic effects. This is borne out in part by the involvement of positively charged arginine residues, capable of participating in bidentate interactions with two negatively charged residues in

these ion-pair networks. The majority of these networks occur through the inter-subunit interface of the hexameric enzymes. In *P. furiosos* GluDH for example, an extensive 18-residue and 6 residue ion-pair network lies at the subunit interface (Yip *et al*., 1998).

Several attempts have been made to introduce more ion pair networks to increase thermostability in GluDHs. *Thermococcus litoralis* GluDH has a half-life of 0.3h at 104 °C compared to 4.6 h in *P. furiosus*; and a melting temperature of 109 °C compared to 114.5 °C in *P. furiosus* (Britton *et al*., 1999). *T. litoralis* GluDH lacks a 6-residue ion-pair network due to the replacement of glutamate residue 138 with threonine (Fig. 2, Vetriani *et al*., 1998). When this threonine residue in *T. litoralis* is mutated to glutamate, an unexpected lowering of thermostability occurred (half-life less than 0.01h at 104 °C; $T_m = 103.5$ °C). Close comparison of the structure of the *T. litoralis* and *P. furiosus* enzymes, showed that there is a subtle shift in relative positions of the main chain and side chain atoms around this six-residue network.

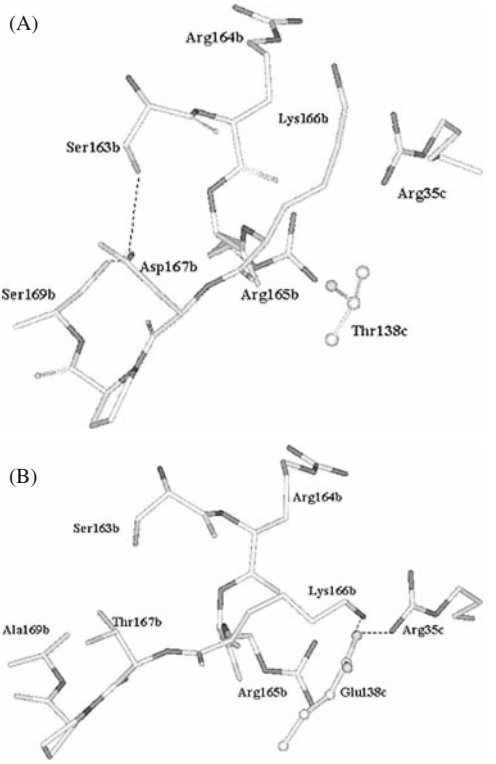


Figure 2. Structures of GluDH from (A) *T. litoralis* (Britton *et al*., 1999) and (B) *P. furiosus*, (Yip *et al*., 1995) showing the contribution of residue 138 in forming stabilizing salt bridges in the *P. furiosus* enzyme. Lower case b and c for each amino acid residue refers to different subunits in the hexameric enzymes

In particular, *T. litoralis* has a nearby aspartate residue (position 167) that forms hydrogen bond interactions with two serine residues (residue 163 and 169). This aspartate is replaced with threonine (T167) in *P. furiosus* GluDH. It was argued that this difference has perturbed the positions of the engineered ion-pair network in *T. litoralis*. Two other mutants are therefore constructed; the single mutant D167T and the double mutant T138E/D167T. The D167T is less stable than the wild-type enzyme but the double mutant possessed a dramatic increase in thermostability (Half life of 1.1h at 104 °C and $T_m = 111.5$ °C). This mutagenesis experiment confirmed the contribution of ion-pair networks in the thermostability of GluDH and also highlighted the importance of other residues, which may not be directly involved in forming the salt-bridges but were important in maintaining proper spatial positions of interacting charged residues. The high sequence identity between *P. furiosus* and *T. litoralis* GluDH (87%) may therefore be a critical element in the successful mutagenesis experiment.

Attempts to engineer increased thermostability in the more distantly related GluDH from *Thermotoga maritima* (55% sequence identity with *P. furiosus*; half-life of 209 min at 85 °C; $T_m = 989.1$ °C) proved more difficult. Two sets of mutagenesis experiments have been performed to introduce ion-pair networks found in *P. furiosus* into the *T. maritima* enzyme. The first experiment involves the construction of a five residue network found in the hinge region, between the two domains of the enzyme subunit (Lebbink *et al*., 1998). This involves the replacement of asparagine 97 in the *T. maritima* enzyme with aspartate (N97D) and glycine 376 with lysine (G376K). The single and double mutants all possess similar melting temperatures as the wild-type but half lives at 85 °C are reduced by about 36%. Since the hinge region is important for the movement of the two domains of the enzyme during catalysis, the mutants displayed temperature dependent differences in kinetic parameters. From 25 °C to 58 °C, V_{max} values in the reverse reaction for L-glutamate production is increased by 1.5-fold in the wild-type enzyme but 2.5-fold in the double mutant. This suggests that "rigidifying" mutations at the hinge region of GluDH may contribute to thermoactivity rather than thermostability.

The second experiment involves the introduction of 16 out of the 18 ion-pair interactions in the dimer interface of *P. furiosus* into the *T. maritima* enzyme (Fig. 3; Lebbink *et al*., 1999). This requires four mutations: S128R, T158E, N117R and S160E. Combinations of mutations were assessed for their contributions to thermostability. Increased thermostability was seen only in the triple mutant (S189R/T158E/N117R) that has a slightly longer half-life at 85°C (240 mins) and higher melting temperature (89.6°C). The lower thermostability of the two single mutants S128R and T158E can be explained by unbalanced charge compensation, since the two residues interact to form a salt bridge. Mutants containing the S160E, including the quadruple mutant, have a large destabilizing effect in the protein, and the reason for this is not clear.

Recently, structure analysis of GluDH from a hyperthermophilic Archaebacteria *Pyrobaculum islandicum* revealed a surprising small number of inter-subunit ionic interactions compared to the *P. furiosus* enzyme (6 versus 54) (Bhuiya *et al*., 2005).

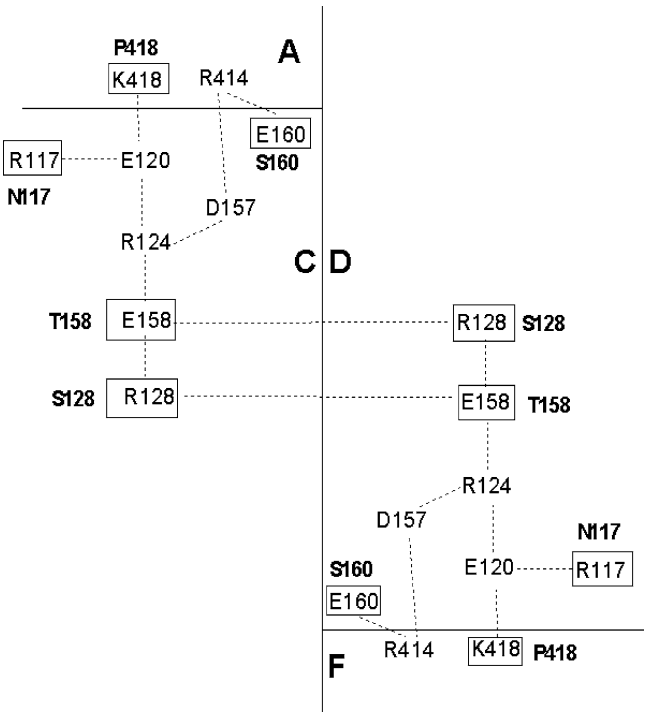


Figure 3. Inter-subunit ion-pair network in *P. furiosus* GluDH. Boxed residues are replaced with other residues in *T. litoralis* (bold beside each box). A, C, D and F refer to the different subunits in the hexameric enzyme. Dotted lines indicate ionic interactions

The total hydrophobic interface between subunits of *P. islandicum* GluDH is however much higher than that of *P. furiosus* (45% versus 20%). The involvement of hydrophobic interactions has also been suggested to contribute to thermostability in *T. maritmia* GluDH (Knapp *et al*, 1997). *T. maritma* is less thermostable than *P. islandicum*, and future comparative analysis of the structure of the two enzymes and site specific mutagenesis studies may shed further light on the contribution of hydrophobic interactions in the stablization of GluDHs.

Thermostable members of other amino acid dehydrogenases have been described (Takada *et al*, 1991; Ohshima *et al*, 1994). For example, LeuDH from *Thermoactinomyces intermedius* has been reported to be stable at 70 °C for 40 mins in 3M NaCl (Ohshima *et al*, 1994). This LeuDH is octameric and the intersubunit contacts differ from the hexameric GluDH mentioned above (Baker *et al*, 1995). Therefore the exact stabilizing interactions in this LeuDH would likely differ from the thermostable GluDHs. The basis for the ability of this LeuDH to remain active at high sodium chloride concentration has also not yet been determined. Molecular modeling of GluDH from a halophile, *Halobacterium salinarium*, showed the presence of a large number of acidic residues on the surface of the enzymes that

may possibly trap a layer of positive charged ions, thus resulting in a formation of hydration shell that prevents protein aggregation under conditions of low water activity (Britton *et al*, 1998). Future structural analysis of LeuDH from *T. intermedius* may yield useful points of comparison between temperature/salt stability within hexameric and octameric members of this family of enzymes.

3.2. Substrate Specificity

The amino acid dehydrogenases are classified according to their substrate specificity (Table 1). In general, GluDHs have narrow substrate specificity. LeuDHs and ValDHs are active towards a variety of aliphatic hydrophobic amino acids. PheDHs have the highest activity with aromatic amino acids, but they also exhibit some activity towards hydrophobic aliphatic amino acids. Sequence identities between PheDHs and LeuDHs/ValDHs are about 50%, while the sequence identities between these groups of enzymes with GluDHs are only about 20%.

In the crystal structure of *C. symbiosum* GluDH, L-glutamate was bound at the base of a deep cleft between the two domains of the enzyme (Stillman *et al*, 1993). The major interactions determining amino acid specificity are formed at the base of this cleft. The 5-carboxyl of the substrate interacts with the side chain of lysine 89 and serine 380, while side chains of valine 377 and alanine 163 form hydrophobic interactions with carbons 4 and 5 of the substrate. Although LeuDHs, ValDHs and PheDHs share low sequence identity with GluDH, sequence alignment and homology modelling shows that the inner core and active site residues in these enzymes are similar to *C. symbiosum* GluDH (Britton *et al*, 1993). Thus, three glycine residues that orient functional groups and determine

Table 1. **Substrate specificity of selected members of GluDH, LeuDH, ValDH and PheDH**. Values for each substrate represent percentages of specific activities relative to the substrate giving the highest activity for each enzyme. ND indicates not determined

Enzyme	GluDH		LeuDH		ValDH		PheDH	
Substrates	Bovine	*C. symbiosum*	*B. cereus*	*B.subilis*	*S. fradiae*	*A. faecalis*	*B. sphaericus*	*B. badius*
Glutamate	100	100	0	0	0	0	0	0
Leucine	1.7	0.026	100	100	25	73	1.3	3.0
Norleucine	1.6	ND	6	40	52	16	3.9	ND
Isoleucine	0.95	0.014	61	65	29	72	0.45	0.20
Valine	1.6	ND	61	92	100	100	1.4	4.0
Norvaline	17	0.052	28	7	98	44	1.3	ND
Methionine	0.82	ND	3	0.18	ND	2	3.0	8.0
Phenylalanine	0	0	0	0	0	0	100	100
Tyrosine	0	0	0	0	0	0	72	9.0
References	Smith *et al*., 1975	Syed, 1987	Schütte *et al*., 1985	Livesey and Lund, 1988	Vandura *et al*., 1988	Ohshima and Soda, 1993	Asano *et al*., 1987a	Asano *et al*., 1987b

the overall shape of GluDH active site are conserved in LeuDH and PheDH. The catalytic lysine (K125) and aspartate (D165) residues are also conserved. However, K89 and S380 are replaced with hydrophobic side chain residues, leucine and valine, respectively, in the modelled structures of LeuDH and PheDH, thus enabling these enzymes to bind hydrophobic substrates. In PheDH, leucine 377 replaces valine in GluDH and LeuDH, to push the substrate to the side of the pocket where extra space is created by a glycine residue in position 163 instead of an alanine found in LeuDH and GluDH. The smaller glycine residue at position 163 allows large aromatic substrate to bind within the active site of PheDH. Site-specific mutagenesis studies have generally supported the predictions of this model.

Glycine 124 (corresponding to position 163 in GluDH) and leucine 307 (corresponding to position 377 in GluDH) in PheDH from *B. sphaericus* have been replaced with glycine and valine, respectively, to create a substrate binding pocket resembling that of LeuDH (Seah *et al*, 1995; Seah *et al*, 2003). Single and double mutants all show a reduction in the specificity for aromatic amino acids while specificity for aliphatic amino acids are increased. The most dramatic result was for the double mutant, which showed a 40-fold increase for the substrate norvaline and a 100-fold increase for the ketoacid in the reverse direction. However, the PheDH mutants are still able to utilize aromatic substrates, albeit poorly. Other subtle changes in the substrate binding pocket, due to differences in as yet unidentified residues, must therefore exist to account for differences in the PheDH mutants and LeuDH.

Site-specific mutagenesis has also been performed on *C. symbiosum* GluDH, by replacing the residue K89 and S380 with the corresponding residues found in LeuDH. Activity towards glutamate is reduced by four orders of magnitude while activities with hydrophobic amino acids are increased in the single mutant K89L. S380V and the double mutant K89L/S380V are however inactive towards glutamate and other aliphatic hydrophobic amino acids tested (Wang *et al*, 1995). Crystallographic structural analysis of the double mutant showed a disordering of loops connecting the two domains in the enzyme subunit and a steric clash between the introduced valine 380 and threonine 193 (Baker *et al*, 1997). It is believed that the different quaternary structures between GluDH and LeuDH (hexameric and octameric) created subtle differences in main chain atoms that resulted in the inability of GluDH to accommodate a valine side chain at position 380. Changing the residue in position 380 of GluDH to alanine alleviated the problems associated with steric hindrance. Hence, the single mutant S380A resulted in the reduction in specificity for L-glutamate by 286-fold, although specific activity towards aliphatic hydrophobic amino acid is still similar to the wild-type enzyme (Wang *et al*, 2001). It was then discovered that the amino acid residue at position 163 within the active site played a role in the recognition of long chain hydrophobic aliphatic amino acids. Thus, the A163G GluDH mutant displayed a 1055-fold reduction in specificity for L-glutamate and 167-fold increase in specificity for L-norleucine at pH 8.0. The mutant also possessed activity with L-methionine at

the same pH, while the wild-type enzyme has no measurable activity with this substrate. Combining the three mutations K89L/S380A/A163G resulted in a further increase in specificity towards L-norleucine and L-methionine by 11.7 fold and 3.3-fold, respectively, compared to the A163G single mutant.

The basis for the differential specificity for L-phenylalanine and L-tyrosine in PheDHs has also been explored. The enzyme from *B. sphaericus* for example displayed similar specificities towards these two aromatic substrates. Specific activity for L-tyrosine in the PheDH of *B. badius* and *S. ureae*, on the other hand, are only 9% and 5.4% the activity with L-phenylalanine. Although the *B. sphaericus* enzyme has favourable thermostable properties, its high specificity towards tyrosine makes it unsuitable for use in the measurement of serum L-phenylalanine levels for diagnosis of phenylketonuria, due to interference from L-tyrosine. Sequence alignment of *B. sphaericus* PheDH with those from *B. badius*, *S. ureae* and *T. intermedius*, showed that there are 10 residues in *B. sphaericus* that are replaced with chemically distinct residues in the other three enzymes (Seah *et al*, 2002). Using the structure of *C. symbiosum* GluDH as a guide, it appears that amide group of asparagine 145 in the *B. sphaericus* enzyme (T193 in *C. symbiosum* GluDH), may provide favourable interactions with the hydroxyl group of the tyrosine substrate. This residue is replaced with valine or alanine in the other PheDHs. Four mutants of *B. sphaericus* PheDH were constructed (N145A, N145L, N145I and N145V) and they all show dramatic increase in discrimination between L-phenylalanine and L-tyrosine, of up to 58-fold seen in the N145I mutant. Surprisingly, this is mainly due to increased specificity for L-phenyalanine, due to a lowered K_m value for this substrate. Kinetic analysis of the wild-type *B. badius* enzyme also showed qualitatively similar results *i.e.* the higher specificity for L-phenylalanine compared to L-tyrosine is due to a lower K_m value for L-phenylalanine. The results can be explained by a more hydrophobic binding pocket in the N145 mutants. Accordingly, specificity for aliphatic hydrophobic amino acid substrates was found to also increase in the N145 mutant enzymes.

Recently, some variations in the sequences of the core residues in the substrate binding sites have been found in several members of the amino acid dehydrogenases. For example, the tetrameric GluDH from *Streptomyces clavuligerus* appear to have a glycine instead of an alanine corresponding to position 163 of *C. symbiosum* GluDH (Miñambres *et al*, 2000). In addition, an arginine residue is found to replace the lysine residue (K89) that interacts with the 5-carboxyl group of the glutamate substrate. *Rhodococcus* sp. M4 PheDH, on the other hand, has an alanine instead of leucine found in all other PheDHs corresponding position 377 of *C. symbiosum* GluDH (Seah *et al*, 2002). Future mutagenesis experiments may shed further light on how these differences contribute to the overall substrate specificity of the enzymes.

Modelling and site-specific mutagenesis studies showed that substrate specificity can be modified in this group of enzymes with relatively few changes in amino acid residues at the active site. As seen from the structures of LeuDH and GluDH, quaternary structures may also have subtle effects on main chain atoms

that ultimately influence substrate specificity. Similarly, a second shell of residues outside the active site may be responsible for the differential specificity for leucine and valine in LeuDHs and ValDHs (Turnbull *et al*, 1997). This indicated that fine-tuning of substrate specificity in this group of enzymes may prove more challenging then the initial mutagenesis studies reported above.

4. BIOTECHNOLOGICAL APPLICATIONS

4.1. Biosynthesis of Chiral Amino Acids

The stereospecific reaction catalyzed by the amino acid dehydrogenases makes them ideal biocatalysts for the synthesis of chiral amino acids for dietary supplements or pharmaceuticals. L-phenylalanine, for example, is a component of the dipeptide aspartame, used as a sugar substitute and as a sweetener in the food and beverage industry. Most natural amino acids, such as L-glutamate, can however be produced in higher yields from the traditional fermentation route, and therefore the amino acid dehydrogenase are more economically useful for the synthesis of unnatural amino acids for the pharamceutical industry. Examples of chiral amino acids used in the drug industry include D-phenylglycine for semisynthetic penicillin, L-homophenylalanine as a component of antihypertensive drugs and L-*tert*-leucine as component of HIV protease and matrix metalloprotease inhibitors.

Due to thermodynamic and economic reasons, the use of the amino acid dehydrogenases (and other dehydrogenases) in the industry requires some type of coenzyme recycling. The most common strategy is to use formate dehydrogenase, which converts the NADH back to NAD^+, at the same time producing carbon dioxide from formate, which can be easily removed from the reactor (Hummel and Kula, 1989). In practice, ammonium formate is used, thus providing the substrate for formate dehydrogenase as well as the ammonium for the amino acid dehydrogenase. Alternatively, glucose dehydrogenase can also be used to recycle the coenzyme, with the formation of glucuronic acid from glucose (Patel, 2001).

Typically, purified amino acid dehydrogenase and an the coenzyme immobilized on resins or PEG can be used in a membrane reactor, and the products can be filtered through a 5000 cut-off membrane for downstream processing (Hummel and Kula, 1989). There has also been some studies using intact recombinant bacteria overexpressing the amino acid dehydrogenase for synthesis of chiral amino acids (Galkin *et al*, 1997). Although artificial coenzyme recycling is not required and enzymes may be more stable intracellularly, higher cost for downstream processing to separate the desired amino acid from complex culture media, as well as side reactions resulting from bacterial aminotransferase activities are some disadvantages of this approach.

The precursor for the amino acid can be provided in the reactor as the keto acid. In cases, where the synthesis of keto acid is difficult, the racemic amino acid can also be used as precursors. This can be illustrated in the synthesis of L-6-hydroxynorleucine, an intermediate for synthesis of the hypertensive drug

Omapatrilat (Patel, 2001). DL-6-hydroxynorleucine can be produced by acid hydrolysis of the commerically available 5-(4-hydroxybutyl)hydantoin. To convert D-6-hydroxynorleucine to L-6-hydroxynorleucine two enzymes are used; D-amino acid oxidase which preferentially converts the D-6-hydroxynorleucine to the corresponding keto acid. This can then be acted on by bovine GluDH to form L-6-hydroxynorleucine. Catalase is used to remove the hydrogen peroxide formed from the D-amino acid oxidase reaction, and glucose dehydrogenase is used to regenerate the NAD^+ cofactor. This process produces complete conversion of the racemic mixture to the L-form with enantiomeric excess of 99%.

The use of L-amino acid dehydrogenase is not restricted to the production of L-amino acids. For example D-*tert*-leucine can be produced from a chemically synthesized racemic mixture of DL-*tert*-leucine (Hummel *et al*, 2003). In this process LeuDH is used in the reverse reaction to remove L-*tert*-leucine by transforming it to the coresponding keto acid. To drive the reaction to completion, and to regenerate the cofactor, NADH oxidase from *Lactobacillus brevis* is used. The D-*tert* leucine left in the mixture has an enantimeric excess of 99%. Of course, the keto acid must be separated from the D-*tert*-leucine, and it is not clear of this can be easily achieved at a reasonable cost.

The applications mentioned above have thus far relied on the use of naturally occurring amino acid dehydrogenases. There is considerable scope to extending the application by using engineered enzymes with unique substrate specificity. For example the N145 mutants of *B. sphaericus* PheDH mentioned in section 1.2.2 can be applied to the synthesis of phenylalanine analogues with hydrophobic substituents at the 4-position of the aromatic ring (Busca *et al*, 2004).

4.2. Amino Acid Dehydrogenases as Diagnostic Reagents

Amino acid dehydrogenase can be used for quantitative measurements of amino acids, ammonia or urea in the clinical setting. For the determination of urea, stoichiometric conversion to ammonia is first achieved through the use of the enzyme urease. In the presence of the keto acid substrate and the amino acid dehydrogenase, the oxidation of NAD(P)H in the reductive deamination direction can be correlated with the concentration of ammonia formed. Traditionally, bovine glutamate dehydrogenase is used for this purpose, but it has also been demonstrated that other amino acid dehydrogenases, such as leucine dehydrogenase can also be used in the same way (Roch-Ramel, 1967; Morishita *et al*, 1997).

L-phenylalanine detection in human serum for diagnosis of the metabolic disease phenylketonuria, is one of the most successful example of the clinical application of amino acid dehydrogenase. Phenylketonuria is a genetic disease caused by phenylalanine hydroxylase deficiency. Individuals with this condition have heightened levels of L-phenylalanine in the blood, of up to 2.4 mM, compared to about 60 μM in normal individuals (Hummel *et al*, 1988). Untreated individuals with PKU suffer from mental retardation. This can be prevented if diagnosis is early and individuals given a strict diet with low levels of L-phenylalanine. As such, most

developed countries screen for this disease in neonates. The traditional method is the Guthrie bacteriological inhibition assay (Guthrie and Susi, 1963). It is based on the fact that L-phenylalanine prevents the growth inhibition of *Bacillus subtilis* by the phenylalanine analogue, ß-theinylalanine. However, this method is only semi-quantitative, labour intensive and it is not suitable for infants undergoing antibiotic therapy. In contrast, an enzymatic assay is rapid, quantitative and can be adapted to a high throughput microplate system. There are two ways to measure phenyalanine levels using PheDH. One involves the measurement of rate of reaction that can be correlated with the absolute concentration of phenyalanine in the sample. The other is the end point assay, where the reaction is allowed to go to completion and the final concentration of coenzyme reduced, which is directly related to the amount of L-pehnylalanine present, can be calculated using its extinction coefficient. Both methods measures the NADH produced which can be determined directly spectrophotometrically at 340nm or fluorimetrically (excitation 340nm, and measurement at 450nm). It is also possible to couple the NADH production with the reduction of a redox compound, that will produce a detectable color change. For example, a method involving reduction of Co^{3+} to Co^{2+} by NADH that is mediated by the electron carrier 1-methoxy-5-methylphenazium methylsulfate has been developed (Naruse *et al*, 1992). Co^{2+} then reacts with 2-(5-bromo-2-pyridazo)-5 N-propyl-N-sulfopropylamino) phenol to produce an intensely coloured species at 590 nm. Alternatively, a biosensor which couples the L-phenylalanine dependent NADH production by PheDH to salicylate hydroxylase and tyrosinase has been developed and found to produce a linear response from 20–150 μM of L-phenyalanine (Huang *et al*, 1998). The advantage of this biosensor is the reusability of the system that can ultimately reduce cost.

Choice of enzyme system for the quantitative determination of amino acids requires a careful consideration of concentration of analyte and the kinetic parameters for the enzyme. If the rate of reaction is measured, in order to maintain linearity between rates and analyte concentrations, the concentration of the analyte must be small relative to the K_m of the enzyme ($[S] < K_m$). This can be achieved by diluting the sample. The low K_m for L-phenylalanine in PheDH is also important for measuring phenylalanine levels in human serum by the end point assay without interference from L-tyrosine. For example, the *B. badius* enzyme used in the kit provided by Sapporo (Japan) has similar k_{cat} values for L-phenylalanine and L-tyrosine ($39 s^{-1}$ and $34 s^{-1}$). However, levels of these two amino acids in the human serum are about 0.06 mM in normal individuals. At these concentrations, the *B. badius* enzyme is acting at about 41% of its maximum rate for L-phenylalanine (K_m 0.087 mM), but at a rate of only 0.7% of its maximum rate for L-tyrosine (K_m 8.6 mM). Individuals with phenylketonuria have heightened levels of L-phenylalanine and therefore background activity with L-tyrosine is negligible.

The use of LeuDH for detection of heightened levels of branched chain amino acids in individuals with maple syrup disease has been described (Wendel *et al*, 1993). Applications for detection of other in-born errors of amino acid metabolism,

such as homocystinuria, is currently limited by the lack of natural enzymes which are sufficiently specific for the particular amino acid.

5. CONCLUSION

Structure and function studies of the amino acid dehydrogenases over the past decade have allowed predictions to be made and tested regarding the molecular basis of substrate specificity and stability in these enzymes. There is still much to be learned about this family of enzymes, such as how allosteric regulation is mediated in the mammalian GluDHs and the basis of the differential coenezyme specificity in various members. Although this chapter has highlighted the success of rational engineering approaches, modern methods of directed evolution, involving random mutagenesis and DNA shuffling, coupled with a robust screening method, may provide a useful alternative strategy to create enzymes with desired properties. This will extend the utility of this important family of enzymes for a wider range of biotechnological processes.

REFERENCES

Asano, Y., Nakazawa, A. and Endo, K. (1987a). Novel phenylalanine dehydrogenases from Sporosarcina ureae and *Bacillus sphaericus*. Purification and characterization. J. Biol. Chem. *262*, 10346–10354.

Asano, Y., Nakazawa, A., Endo, K., Hibino, Y., Ohmori, M., Numao, N. and Kondo, K. (1987b). Phenylalanine dehydrogenase of Bacillus badius. Purification, characterization and gene cloning. Eur. J. Biochem. *168*, 153–159.

Baker, P.J., Turnbull, A.P., Sedelnikova, S.E., Stillman, T.J. and Rice, D.W. (1995). A role for quaternary structure in the substrate specificity of leucine dehydrogenase. Structure. *3*, 693–705.

Baker, P.J., Waugh, M.L., Wang, X.G., Stillman, T.J., Turnbull, A.P., Engel, P.C. and Rice, D.W. (1997). Determinants of substrate specificity in the superfamily of amino acid dehydrogenases. Biochemistry *36*, 16109–16115.

Bhuiya, M.W., Sakuraba, H., Ohshima, T., Imagawa, T., Katunuma, N. and Tsuge, H. (2005). The first crystal structure of hyperthermostable NAD-dependent glutamate dehydrogenase from Pyrobaculum islandicum. 2005. J Mol Biol. *345*, 325–337.

Britton, K.L., Baker, P.J., Engel, P.C., Rice, D.W. and Stillman, T.J. (1993). Evolution of substrate diversity in the superfamily of amino acid dehydrogenases. Prospects for rational chiral synthesis. J. Mol. Biol. *234*, 938–945.

Britton, K.L., Stillman, T.J., Yip, K.S., Forterre, P., Engel, P.C. and Rice, D.W. (1998). Insights into the molecular basis of salt tolerance from the study of glutamate dehydrogenase from *Halobacterium salinarum*. J. Biol. Chem. *273*, 9023–9030.

Britton, K.L., Yip, K.S., Sedelnikova, S.E., Stillman, T.J., Adams, M.W., Ma, K., Maeder, D.L., Robb, F.T., Tolliday, N., Vetriani, C., Rice, D.W. and Baker, P.J. (1999). Structure determination of the glutamate dehydrogenase from the hyperthermophile *Thermococcus litoralis* and its comparison with that from *Pyrococcus furiosus*. J. Mol. Biol. *293*, 1121–1132.

Brunhuber, N.M. and Blanchard, J.S. (1994). The biochemistry and enzymology of amino acid dehydrogenases. Crit. Rev. Biochem. Mol. Biol. *29*, 415–467.

Brunhuber, N.M., Thoden, J.B., Blanchard, J.S. and Vanhooke, J.L. (2000). *Rhodococcus* L-phenylalanine dehydrogenase: kinetics, mechanism, and structural basis for catalytic specificity. Biochemistry. *39*, 9174–9187.

Busca, P., Paradisi, F., Moynihan, E., Maguirem, A.R. and Engel, P.C. (2004). Enantioselective synthesis of non-natural amino acids using phenylalanine dehydrogenases modified by site-directed mutagenesis. Org. Biomol. Chem. *2*, 2684–2691.

Dean, J.L., Wang, X.G., Teller, J.K., Waugh, M.L., Britton, K.L., Baker, P.J., Stillman, T.J., Martin, S.R., Rice, D.W. and Engel, P.C. (1994). The catalytic role of aspartate in the active site of glutamate dehydrogenase. Biochem. J. *301*, 13–16.

Elcock, A.H. (1998). The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. J. Mol. Biol. *284*, 489–502.

Galkin, A., Kulakova, L., Yoshimura, T., Soda, K. and Esaki, N. (1997). Synthesis of optically active amino acids from alpha-keto acids with *Escherichia coli* cells expressing heterologous genes. Appl. Environ. Microbiol. *63*, 4651–4656.

Guthrie, R. and Susi, A. 1963. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. Pediatrics. *32*, 338–343.

Hayden, B.M., Dean, J.L., Martin, S.R. and Engel, P.C. (1999). Chemical rescue of the catalytically disabled clostridial glutamate dehydrogenase mutant D165S by fluoride ion. Biochem. J. *340*, 555–560.

Huang, T., Warsinkem, A., Kuwana, T. and Scheller, F.W. (1998). Determination of L-phenylalanine based on an NADH-detecting biosensor. Anal. Chem. *70*, 991–997.

Hudson, R.C. and Daniel, R.M. (1993) L-glutamate dehydrogenases: distribution, properties and mechanism. Comp. Biochem. Physiol. *106B*, 767–792.

Hummel, W., Schutte, H. and Kula, M.R. (1988). Enzymatic determination of L-phenylalanine and phenylpyruvate with L-phenylalanine dehydrogenase. Anal. Biochem. *170*, 397–401.

Hummel, W. and Kula, M.-R. (1989). Dehydrogenases for the synthesis of chiral compounds. Eur. J. Biochem. *184*, 1–13.

Hummel, W., Kuzu, M. and Geueke, B. (2003). An efficient and selective enzymatic oxidation system for the synthesis of enantiomerically pure *D-tert*-leucine. Org. Lett. *5*, 3649–3650.

Hyun, C.G., Kim, S.S., Park, K.H. and Suh, J.W. (2000). Valine dehydrogenase from *Streptomyces albus*: gene cloning, heterologous expression and identification of active site by site-directed mutagenesis. FEMS Microbiol. Lett. *182*, 29–34.

Kataoka, K., Tanizawa, K., Fukui, T., Ueno, H., Yoshimura, T., Esaki, N. and Soda, K. (1994). Identification of active site lysyl residues of phenylalanine dehydrogenase by chemical modification with methyl acetyl phosphate combined with site-directed mutagenesis. J. Biochem. *116*, 1370–1376.

Knapp, S., de Vos, W.M., Rice, D. and Ladenstein, R. (1997). Crystal structure of glutamate dehydrogenase from the hyperthermophilic eubacterium *Thermotoga maritima* at 3.0 A resolution. J. Mol. Biol. *267*, 916–932.

Lebbink, J.H., Knapp, S., van der Oost, J., Rice, D., Ladenstein, R. and de Vos, W.M. (1998). Engineering activity and stability of *Thermotoga maritima* glutamate dehydrogenase. I. Introduction of a six-residue ion-pair network in the hinge region. J. Mol. Biol. *280*, 287–296.

Lebbink, J.H., Knapp, S., van der Oost, J., Rice, D., Ladenstein, R. and de Vos, W.M. (1999). Engineering activity and stability of *Thermotoga maritima* glutamate dehydrogenase. II: construction of a 16-residue ion-pair network at the subunit interface. J. Mol. Biol. *289*, 357–369.

Leiser, A., Birch, A. and Robinson, J.A. (1996). Cloning, sequencing, overexpression in *Escherichia coli*, and inactivation of the valine dehydrogenase gene in the polyether antibiotic producer *Streptomyces cinnamonensis*. Gene *177*, 217–222.

Livesey, G. and Lund, P. (1988). Determination of branched-chain amino and keto acids with leucine dehydrogenase. Methods Enzymol. *166*, 3–10.

Massey, L.K., Sokatch, J.R. and Conrad, R.S. (1976). Branched-chain amino acid catabolism in bacteria. Bacteriol. Rev. *40*, 42–54.

Miñambres, B., Olivera, E.R., Jensen, R.A. and Luengo, J.M. (2000). A new class of glutamate dehydrogenases (GDH). Biochemical and genetic characterization of the first member, the AMP-requiring NAD-specific GDH of *Streptomyces clavuligerus*. J. Biol. Chem. *275*, 39529–39542.

Morishita, Y., Nakane, K., Fukatsu, T., Nakashima, N., Tsuji, K., Soya, Y., Yoneda, K., Asano, S. and Kawamura, Y. (1997). Kinetic assay of serum and urine for urea with use of urease and leucine dehydrogenase. Clin. Chem. *43*, 1932–1936.

Naruse, H., Ohashi, Y.Y., Tsujii, A., Maeda, M., Nakamura, K., Fujii, T., Yamaguchi, A., Matsumoto, M. and Shibata, M. (1992). A method of PKU screening using phenylalanine dehydrogenase and microplate system. Screening *1*, 63–66.

Ohshima, T., Nishida, N., Bakthavatsalam, S., Kataoka, K., Takada, H., Yoshimura, T., Esaki, N. and Soda K. (1994). The purification, characterization, cloning and sequencing of the gene for a halostable and thermostable leucine dehydrogenase from *Thermoactinomyces intermedius*. Eur. J. Biochem. *222*, 305–312.

Oikawa, T., Yamanaka, K., Kazuoka, T., Kanzawa, N. and Soda, K. (2001). Psychrophilic valine dehydrogenase of the antarctic psychrophile, *Cytophaga sp*. KUC-1: purification, molecular characterization and expression. Eur. J. Biochem. *268*, 4375–4383.

Patel, R.N. (2001). Enzymatic synthesis of chiral intermediates for Omapatrilat, an antihypertensive drug. Biomol. Eng. *17*, 167–182.

Peterson, P.E. and Smith, T.J. (1999). The structure of bovine glutamate dehydrogenase provides insights into the mechanism of allostery. Structure Fold. Des. *7*, 769–782.

Rice, D.W., Hornby, D.P. and Engel, P.C. (1985). Crystallization of an NAD^+-dependent glutamate dehydrogenase from *Clostridium symbiosum*. J. Mol. Biol. *181*, 147–149.

Rife, J.E. and Cleland, W.W. (1980). Determination of the chemical mechanism of glutamate dehydrogenase from pH studies. Biochemistry *19*, 2328–2333.

Roch-Ramel, F. (1967). An enzymic and fluorophotometric method for estimating urea concentrations in nanoliter specimens. Anal. Biochem. *21*, 372–381.

Seah, S.Y.K., Britton, K.L., Baker, P.J., Rice, D.W., Asano, Y. and Engel, P.C. (1995). Alteration in relative activities of phenylalanine dehydrogenase towards different substrates by site-directed mutagenesis. FEBS Lett. *370*, 93–96.

Seah, S.Y.K., Britton, K.L., Rice, D.W., Asano, Y. and Engel, P.C. (2002). Single amino acid substitution in *Bacillus sphaericus* phenylalanine dehydrogenase dramatically increases its discrimination between phenylalanine and tyrosine substrates. Biochemistry *41*, 11390–11397.

Seah, S.Y.K., Britton, K.L., Rice, D.W., Asano, Y. and Engel, P.C. (2003). Kinetic analysis of phenylalanine dehydrogenase mutants designed for aliphatic amino acid dehydrogenase activity with guidance from homology-based modelling. Eur. J. Biochem. *270*, 4628–4634.

Sekimoto, T., Matsuyama, T., Fukui, T. and Tanizawa, K. (1993). Evidence for lysine 80 as general base catalyst of leucine dehydrogenase. J. Biol. Chem. *268*, 27039–27045.

Schütte, H., Hummel, W., Tsai, H. and Kula, M.R. (1985). L-leucine dehydrogenase from *Bacillus cereus* - production, large-scale purification and protein characterization. Appl. Microbiol. Biotechnol. *22*, 306–317.

Smith, E.L., Austen, B.M., Blumenthal, K.M. and Nyc, J.F. (1975). In The Enzymes (Boyer, PD ed.) VolXI, pp. 293–367, Academic Press, New York.

Smith, T.J., Schmidt, T., Fang, J., Wu, J., Siuzdak, G. and Stanley, C.A. (2002). The structure of apo human glutamate dehydrogenase details subunit communication and allostery. J. Mol. Biol. *318*, 765–777.

Stillman, T.J., Baker, P.J., Britton, K.L. and Rice, D.W. (1993). Conformational flexibility in glutamate dehydrogenase. Role of water in substrate recognition and catalysis. J. Mol. Biol. *234*, 1131–1139.

Syed, S.E.,-H. (1987). Beef liver glutamate dehydrogenase: studies of substrate specificity and relationship between the catalytic sites. Ph.D. Thesis. University of Sheffield, UK.

Takada, H., Yoshimura, T., Ohshima, T., Esaki, N. and Soda, K. (1991). Thermostable phenylalanine dehydrogenase of *Thermoactinomyces intermedius*: cloning, expression, and sequencing of its gene. J. Biochem. *109*, 371–376.

Tang, L., Zhang, Y.H. and Hutchinson, C.R. (1994). Amino acid catabolism and antibiotic synthesis: valine is a source of precursors for macrolide biosynthesis in *Streptomyces ambofaciens* and *Streptomyces fradiae*. J. Bacteriol. *176*, 6107–6119.

Turnbull, A.P., Baker, P.J. and Rice, D.W. (1997). Analysis of the quaternary structure, substrate specificity, and catalytic mechanism of valine dehydrogenase. J. Biol. Chem. *272*, 25105–25111.

Vancura, A., Vancurova, I., Volc, J., Fussey, S.P., Flieger, M., Neuzil, J., Marsalek, J. and Behal, V. (1988). Valine dehydrogenase from *Streptomyces fradiae*: purification and properties. J. Gen. Microbiol. *134*, 3213–3219.

Vanhooke, J.L., Thoden, J.B., Brunhuber, N.M., Blanchard, J.S. and Holden, H.M. (1999). Phenylalanine dehydrogenase from *Rhodococcus sp*. M4: high-resolution X-ray analyses of inhibitory ternary complexes reveal key features in the oxidative deamination mechanism. Biochemistry. *38*, 2326–2339.

Vetriani, C., Maeder, D.L., Tolliday, N., Yip, K.S., Stillman, T.J., Britton, K.L., Rice, D.W., Klump, H.H. and Robb, F.T. (1998). Protein thermostability above 100 degrees C: a key role for ionic interactions. Proc. Natl. Acad. Sci. U.S.A. *95*, 12300–12305.

Wang, X.G., Britton, K.L., Baker, P.J., Martin, S., Rice, D.W. and Engel, P.C. (1995). Alteration of the amino acid substrate specificity of clostridial glutamate dehydrogenase by site-directed mutagenesis of an active-site lysine residue. Protein Eng. *8*, 147–152.

Wendel, U., Gonzales, J. and Hummel, W. (1993). Neonatal screening for maple syrup urine disease by an enzyme-mediated colorimetric method. Clin. Chim. Acta. *219*, 105–111.

Werner, C., Stubbs, M.T., Krauth-Siegel, R.L. and Klebe, G. (2005). The crystal structure of *Plasmodium falciparum* glutamate dehydrogenase, a putative target for novel antimalarial drugs. J. Mol. Biol. *349*, 597–607.

Yip, K.S., Britton, K.L., Stillman, T.J., Lebbink, J., de Vos, W.M., Robb, F.T., Vetriani, C., Maeder, D. and Rice, D.W. (1998). Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. Eur. J. Biochem. *255*, 336–346.

Yip, K.S., Stillman, T.J., Britton, K.L., Artymiuk, P.J., Baker, P.J., Sedelnikova, S.E., Engel, P.C., Pasquo, A., Chiaraluce, R. and Consalvi, V. (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. Structure. *3*, 1147–1158.

CHAPTER 29

PHYTASE: SOURCE, STRUCTURE AND APPLICATION

XIN GEN LEI[1]*, JESUS M. PORRES[2], EDWARD J. MULLANEY[3] AND HENRIK BRINCH-PEDERSEN[4]

[1]*Department of Animal Science, Cornell University, Ithaca, New York, USA*
[2]*Departamento de Fisiología, Universidad de Granada, Granada, Spain*
[3]*SRRC-ARS-USDA, New Orleans, USA*
[4]*Danish Institute of Agricultural Sciences, Department of Genetics and Biotechnology, Slagelse, Denmark*
**xl20@cornell.edu*

1. INTRODUCTION

Phytases have been one of the focal enzymes for nutrition, environmental protection, and human health during the past two decades. These enzymes sequentially cleave orthophosphate groups from the inositol core of phytate, the major chemical form of phosphorus in plants. Various phytases have been isolated from plants and microbes, and can be grouped based on their pH optima (alkaline or acid phytases), catalytic mechanisms (histidine acid phosphatases, ß-propeller phytase, cysteine phosphatases or purple acid phosphatases), or stereospecificity of phytate hydrolysis (3- or 6-phytases). Recent phytase research has been driven by the urgent need for improving utilization of phytate-phosphorus in diets for simple-stomached animals to reduce their manure phosphorus excretion to environment. However, potential applications of phytases may extend to release dietary phytate-bound minerals for human nutrition and to develop special inositol phosphates for human health.

2. SOURCES OF PHYTASE

2.1. Microbes

Phytases have been isolated from fungi, yeast, bacteria, and protozoa. Most these enzymes belong to the histidine acid phosphatase or alkaline phytase sub-families, and exhibit considerable variations in kinetics, stereospecificities, and biochemical properties. Several microbial phytases have been commercialized as animal feed supplements.

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 505–529

2.1.1. Fungal and yeast phytases

Usually classified as 3-phytases, most of phytases isolated from fungi and yeast are histidine acid phosphatases, glycosylated, and active for a wide variety of substrates (Wyss *et al*., 1999a). *Aspergillus niger* PhyA was the first well-characterized and commercialized phytase. Encoded by a 1.4 kb DNA fragment, this enzyme is a monomer with an approximate molecular weight of 80 kDa, a bi-hump pH profile with two optimal pH at 2.5 and 5.0–5.5, an optimal temperature at 55–60 °C, and high affinity for phytic acid (Han *et al*., 1999). *Aspergillus fumigatus* phytase shares a 66% sequence similarity with *A. niger* PhyA phytase, but displays better thermo-tolerance (Pasamontes *et al*., 1997a; Wyss *et al*., 1998). Its thermo-tolerance was related to a great efficiency of refolding after heat denaturation, and can be modulated by specificity of the buffers used in the heat treatment (Rodriguez *et al*., 2000a). The enzyme has a broad range of pH, and is highly active against inositol phosphates with low degree of phosphorylation (Wyss *et al*., 1999a; Rodriguez *et al*., 2000a). However, its specific activity against phytate is low (Tomschy *et al*., 2000). *Peniophora lycii* PhyA phytase has also been commercialized. It is a 6-phytase with an optimal pH at 4.0–4.5 and optimal temperature at 50–55 °C, and has dimeric conformation (Lassen *et al*., 2001). It seems to be susceptible to thermal treatments and proteases (Simon and Igbasan, 2002) or low pH.

Quan *et al*. (2004) have isolated a low molecular weight (32.6 kDa) phytase from the air-borne fungus *Cladosporium* sp. FP-1. The enzyme is not glycosylated, and has an optimal pH at 3.5 and an optimal temperature at 40 °C. It produces inositol tri-phosphate as the final product. Phytases isolated from thermophilic fungi *Myceliophtora thermophila* and *Talaromyces thermophilus* (Mitchell *et al*., 1997; Pasamontes *et al*., 1997b) exhibit a high degree of sequence homology to other fungal phytases from *A. niger, A. terreus* or *A. fumigatus*. Berka *et al*. (1998) isolated a phytase from the thermophilic fungus *Thermomyces lanuginosus* that demonstrated a better thermostability and catalytic efficiency, and a higher transition temperature than that of the *A. niger* phytase. Chadha *et al*. (2004) reported that phytase produced by the thermophilic fungus *Mucor pusillus* was fairly active in a wide pH range of 3 to 7.8.

From a survey on 738 strains of yeast, Nakamura *et al*. (2000) found significant levels of phytase activity in 35 species, with a wide range of optimal pH and temperature. *Arxula adeninivorans* grew well in media containing phytate as the sole source of phosphate, and secreted phytase that had an optimal pH in the range of 4.5–5.0, and optimal temperature around 75 °C (Sano *et al*., 1999). A significant phytase production has also been reported by Quan *et al*. (2002) from the soil-isolated yeast *Candida Krusei* WZ-001. The isolated phytase contained two different subunits with molecular masses of 116 and 31 kDa, had a glycosylation rate of approximately 35%, and exhibited optimal pH and temperature at 4.6 and 40 °C, respectively. Phytase activity has also been detected in *Pichia anomala* (Vohra and Satyanarayana, 2001), *Saccharomyces cerevisiae* (Türk *et al*., 2000), and *Schwanniomyces castellii* (Segueilha *et al*., 1992). These enzymes were active in the acidic pH range, with an optimal temperature at 60–74 °C.

2.1.2. *Bacterial phytases*

Phytases isolated from bacteria are non-glycosylated histidine acid phosphatases or alkaline phytases with a ß-propeller structure. *Escherichia coli* AppA phytase is a periplasmic protein with a molecular mass of approximately 42 kDa. Because of its acidic optimal pH, strong resistance to pepsin hydrolysis and high specific activity for phytic acid (Wyss *et al*, 1999a), *E. coli* AppA phytase is more effective than *A. niger* phytase in releasing phytate-phosphorus in diets for swine and poultry (Augspurger *et al*, 2003). Meanwhile, a novel *E. coli* phytase gene (*appA2*) has been isolated from pig colon, and expressed in *Pichia pastoris* (Rodriguez *et al*, 1999b). *Bacillus* phytases are monomers with a molecular mass of 38–47 kDa, optimal pH in the neutral range and optimal temperature at 55–70 °C (Kerovuo *et al*, 1998).

Yanke *et al*. (1998) identified several phytase-producing anaerobic bacteria in cattle rumen, and found *Selenomonas* species to be the highest producer followed by a strain of *Mitsuokella multiacidus*. Cho *et al* (2003) isolated a phytase enzyme from *Pseudomonas syringae* with molecular mass of 45 kDa, specific activity of 649 U/mg protein, pH optimum at 5.5 and temperature optimum at 40 °C. Kim *et al* (2003) isolated a novel phytase from *Citrobacter braaki* with optimal pH and temperature at 4 and 50 °C, respectively, and higher specificity against phytic acid than other phosphorylated substrates. Phytases have also been isolated from *Obesumbacterium proteus* (Zinin *et al*, 2004), soil bacterium *Klebsiella* spp. ASR1 (Sajidan *et al*, 2004), and several species of the *Bifidobacterium* genera (Haros *et al*, 2005).

2.1.3. *Other micro-organisms*

Freund *et al*. (1992) have reported the existence of phytase in protozoan *Paramecium tetraurelia*. The enzyme appeared to be a hexamer of 240 kDa with optimal pH of 7.0 and no requirements for divalent cations for activity, and was stereospecific in sequentially removing the phosphates at the 6, 5, 4 and 1 positions (Van der Kaay and Van der Haastert, 1995). Cheng (2005) have sequenced a putative β-propeller phytase gene from the psychrophile *Shewanella oneidensis* MR-1 that showed a 30% peptide sequence identity with that of *Bacillus spp* phytase. The importance of this finding lies, based on the presence of cold-active protein-tyrosinase phosphatases isolated from this species, in the potential application of this new phytase to aquaculture.

2.2. Plants

Most of plant phytases are histidine acid phosphatases, with an optimal pH between 4.5–6.0 and optimal temperature between 38–55 °C. However, there are wide variations of plant phytases in kinetics (K_m: 30 to 300 μM; K_{cat}: 43 to 704 s^{-1}; and specific activity: 43 to 636 U/mg protein). Plant phytases in the histidine acid phosphatase family were generally considered as 6-phytase. However, recent data indicated that some of them (Lupin LP11 and LP12) initiated the hydrolysis of the orthophosphate at the D-3 position of the inositol ring (Greiner *et al*, 2002).

Some plant phytases are found to be alkaline phosphatases, or purple acid phosphatases. The phytase from lilly pollen showed an optimal pH of 8 and an optimal temperature of 55 °C (Iog *et al*, 2005). This enzyme was activated by calcium and inactivated by EDTA, and had a narrow substrate specificity, with D-Ins(1,2,3)P3 as the end product. Hegeman (2001) have isolated a phytase gene from germinating soybean that did not share any sequence similarity to histidine acid phosphatase, but a high degree of sequence similarity to purple acid phosphatase that contains a binuclear Fe(III)-Me(II) center. The enzyme displayed optimal pH at 4.5–5.0 and optimal temperature at 58 °C.

Phytase activity has been detected in cereals, legumes, and oil seeds (Viveros *et al*., 2000) or highly consumed fruits and vegetables like the avocado and scallion leaves (Phillippy and Wyatt, 2001). Certain cereal grains (*i.e.* wheat, spelt, rye, barley, triticale) display high levels of phytase activity that can reach more than 5,000 units/kg. Use of these grains and their by-products as a plant phytase source has been tested in animal feeding (Han *et al*, 1997). Industrial or household processing such as germination, fermentation, and soaking can be employed to make good uses of intrinsic phytase activity present in plant foods (Fredlund *et al*, 2003; Porres *et al*, 2003).

2.3. Animal Tissues

Although phytase activity has been detected in tissues of several animal species (Bitar and Reinhold, 1972), there is no complete molecular characterization of any of animal-derived phytases. Many of these enzymes display an optimal pH in the neutral to alkaline range, with K_m for phytate ranging from 0.03 to 2.6 mM. However, phytase detected in brush border vesicles of poultry by Maenz (1998) showed an optimal pH of 5.5–6.0 and phytase in the hybrid stripped bass showed an optimal pH of 3.5–4.5 (Ellestad *et al*, 2002).

Even though phytases have been isolated from the intestinal brush border membrane (Maenz and Classen, 1998; Ellestad *et al*, 2002), their practical importance for improving the availability of dietary phytate-phosphorus to simple-stomached animals could be over-shaded by the very affordable supplementation of exogenous phytase in feed. Phytase activity found in the large intestine or rumen is mainly microbial origin (Wise and Gilburt, 1982; Yanke *et al*, 1998).

3. STRUCTURE

The basic structural features of several phytate degrading enzymes have been established as representatives of previously known classes of phosphatase (Mullaney and Ullah, 2003; Chu *et al*, 2004). In others, X-ray crystallographic studies have confirmed they belong to a class with a novel catalytic mechanism (Ha *et al*, 2000). In both instances, the elucidation of the 3-D molecular structure of different phytate-degrading enzymes has enhanced our understanding of the linkage between the molecular structure of the molecule and it's catalytic function. It is now evident

that different phytases have evolved to supply the unique nutritional requirements found in various forms of plant, animal and microbial life. It also appears that there is a direct link between an enzyme's catalytic domain and specific molecular architectural elements. Thus, while specific structural components are essential, other nonessential parts of the molecule may be altered to adapt the catalytic mechanism to various substrates. The precise number of catalytic mechanisms that nature has evolved to hydrolyze phytate will be determined by future research. At this time, four classes of phosphatase enzymes are known to have representatives that can degrade phytic acid: (1) Histidine Acid Phosphatase (HAP), (2) β-Propellar Phytase (BPP), (3) Cysteine Phosphatase (CP) and (4) Purple Acid Phosphatase (PAP). Each one of these has unique structural features due to their distinct catalytic apparatus that allows them to utilize phytic acid as a substrate in various environments.

3.1. Histidine Acid Phosphatase (HAP)

All members of the HAP share a common catalytic mechanism and common active site motif. The N-terminal active site motif is RHGXRXP and the C-terminal motif is HD (Wodzinski and Ullah, 1996). When the amino acid sequence is properly folded, these components position together to form the catalytic site of this class of phosphatases. These distant sequences converge to form a single catalytic center that initiates a two-step reaction that hydrolyzes phosphomonoesters (Van Etten *et al*., 1991).

HAPs are a large group of acid phosphatases that depending on the species can hydrolyze an array of different substrates. Thus, it is important to realize that not all HAPs can effectively degrade phytate. Phytate is a highly negatively charged substrate and in order for any catalytic mechanism to interact with it, the active site must be able to accommodate this feature. Therefore, in order to facilitate substrate binding, the active site is primarily positively charged at acidic pHs in both prokaryotic and eukaryotic HAPs that effectively hydrolyze phytate. Oh *et al*. (2004) has purposed the term Histidine Acid Phytase (HAPhy) to designate HAPs that can effectively hydrolyze phytate. Both prokaryotic and eukaryotic HAPhys are known. The best-characterized prokaryotic HAPhy is *E. coli* phytase (Greiner *et al*., 1993). A 3-D molecular model of its structure is available (Lim *et al*., 2000). In eukaryotes, HAPhys have been cloned in a number of fungal isolates and in maize (Mullaney *et al*., 2000). The most widely studied fungal phytases are from *A. niger* (Fig. 1) and *A. fumigatus*.

An important factor in determining and maintaining the structure of HAPhys is glycosylation. This process, that adds polysaccharides to proteins, confers stability and assist in the correct folding of the enzyme. All the extra cellular fungal phytases that have been characterized to date are glycoproteins. *A. niger* NRRL 3135 PhyA is heavily glycosylated with ten N-glycosylation sites (Ullah and Dischinger, 1993). Another vital structural component in HAPhys are disulfide bridges that perform an important role in maintaining the proper 3-dimensional structure to allow for catalytic activity in phytase (Wang *et al*., 2004; Mullaney, 2005; Kostrewa *et al*., 1997). All ten cysteine residues present in *A. niger* and *A. fumigatus* PhyA

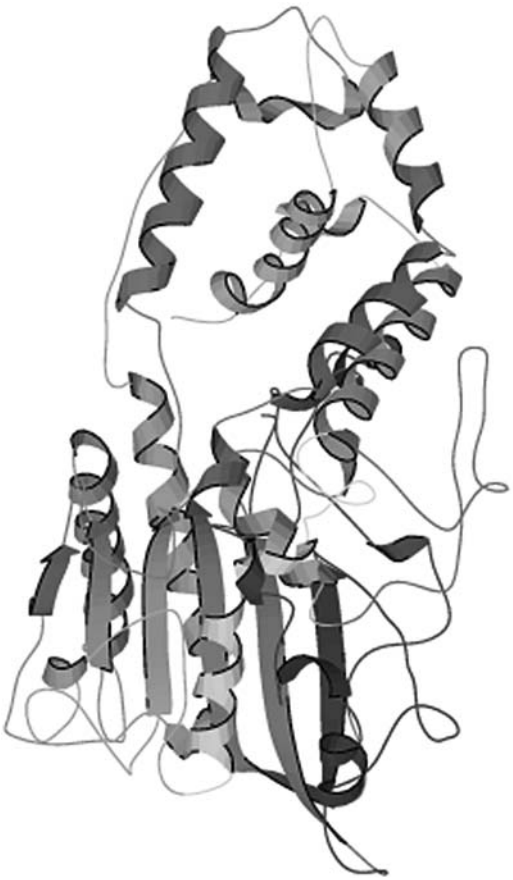


Figure 1. Structure of *Aspergillus (ficuum) niger* NRRL 3135 PhyA (Kostrewa *et al*., 1997), a representative model of histidine acid phosphatases

are involved in the formation of disulfide bridges. All of the eight cysteines in *E. coli* phytase are involved in disulfide bonds (Lim *et al*., 2000). However, in this phytase, significantly enhanced activity was achieved when site-directed mutagenesis was utilized to remove one disulfide bridge (Rodriguez *et al*., 2000b). It was suggested that its removal allowed enhanced domain flexibility and thereby increase the catalytic efficiency of the enzyme.

Structural characterization (Kostrewa *et al*., 1999; Liu *et al*., 2004) and catalytic studies (Wyss *et al*., 1999a) have assigned a vital role to a new site in the enzyme that facilitates its interaction with different substrates. Kostrewa *et al*. (1999a) identified several amino acid residues that constitute a substrate specificity site (SSS) in the *A. niger* PhyA molecule that encircles the enzyme's active site and functions as a "gatekeeper". In the SSS of *A. niger* NRRL 3135 there are two

acidic and four basic amino acid residues: E228, D262, K91, K94, K300 and K301 (Kostrewa *et al*, 1999; Mullaney *et al*, 2000). At pH 2.5 the four basic amino acids; K91, K94, K300 and K301 in the *A. niger* SSS are all positively charged and would attract the phytate molecule. Also, the local electrostatic field of the SSS remains attractive for phytate when the pH is raised to 5.0.

Wyss *et al*. (1999a) has divided all the known microbial HAPhys into two classes based on the substrate specificity. The first class has broad substrate specificity but a low specific activity for phytate, while the second class has narrow substrate specificity and a high specific activity for phytate. It has since been discovered that a correlation exist between amino acid residue 300, a SSS component, and the enzyme's level of specific activity for phytic acid (Mullaney *et al*, 2002). This study also revealed that while amino acid residue 300 varied, residue 301 was strongly conserved as lysine (K). The phytate degrading enzymes cited in Wyss *et al*. (1999a) with high specific activity for phytic have either a basic or acidic amino acid residue at 300, while the phytases with low specific activity have a neutral amino acid at that position. The importance of the lysine residue at that site and the enzyme's high specific activity for phytic acid has been established by site-directed mutagenesis (Mullaney *et al*, 2002).

The importance of the SSS in determining pH optimum and substrate specificity range is evident in a second extracellular *A. niger* phytase, PhyB (Ullah and Cummins, 1987). PhyB has only been reported in the isolates of *A. niger*. PhyB's optimum pH is 2.5 and unlike PhyA displays no ability to hydrolyze phytate at pH 5. While PhyB and PhyA share identical active site characteristics of HAPs, their SSS, are different. Kostrewa *et al* (1999) identified the SSS of PhyB to be composed of only two acidic amino acids, D75 and E272. This explains why PhyB has a different pH profile than PhyA. At pH 5.0 the acidic amino acids in the PhyB SSS would be negatively charged, while at pH 2.5 they would be uncharged. All negatively charged substrates, such as phytate, would therefore be repelled at the higher pH. Because PhyB's SSS has a more neutral electrostatic field, it can accept a broader variety of phosphomonoesters than PhyA. The highly positive electrostatic field of *A. niger's* PhyA SSS is optimized for the binding of a negatively charged substrate, such as phytate.

This evidence indicates the SSS has a significant role in determining how effectively the enzyme can hydrolyze phytate. By occupying positions adjacent to the catalytic domain, the amino acids in the SSS function as "gatekeeper" in determining which substrates can easily pass and interact with the active site residues. Research is also showing that techniques such as site-directed mutagenesis can be employed to alter the composition of the enzyme's SSS and thus, alter both the enzyme's pH profile and substrate range.

3.2. β–Propeller Phytase (BPPhy)

The molecular structure of a thermostable phytase (TS-phy) from *Bacillus amyloliquefaciens* has been identified (Ha *et al*, 2000). This phytate-degrading

enzyme is not a member of the histidine acid phosphatase class of enzymes, but rather represents an entirely new class of enzyme that displays no obvious homology to any known phosphatase class. Unlike the HAPhys, which are members of a well-studied class of enzymes, the *Bacillus* phytases represent an entirely new class of enzymes and exhibit no homology to any known phosphatases (Kim *et al*., 1998; Kerovuo *et al*., 1998; Ha *et al*., 2000). The name β–Propeller Phytase (BPPhy) was adopted for the group after its molecular structure was determined, which consists mainly of β-sheets and resembles a six bladed propeller (Fig. 2) (Ha *et al*., 2000; Shin *et al*., 2001).

The first reported BPPhys were from *Bacillus* and related bacterial species. All these reported enzymes were similar in that they require Ca^{2+} for both catalytic activity and thermostability. The calcium ions are thought to facilitate the binding of phytate by generating a favorable electrostatic environment. Thus, as in the SSS HAPhys the substrate binding domain of the biocatalyst attracts the substrate. Kinetic studies at pH 7.0–8.0 have established that BPPhys can hydrolyze calcium-phytate at that pH range (Oh *et al*., 2001). Two main components are involved in the catalytic mechanism of BPPhys. The "affinity site" attracts the substrate and an adjacent "cleavage site" to hydrolyze the phosphate group (Shin *et al*., 2001). This model explains BPPhy preference for phytate, since it is necessary for two neighboring phosphate groups to occupy both the cleavage and affinity sites. The enzyme prefers hydrolysis of every second phosphate. This explains why this enzyme alternately removes phosphate groups with the end product being *myo*-inositol triphosphate. Degradation of phytate to IP_3 occurs rapidly, but further hydrolysis is not favored because a neighboring phosphate group is lacking.

Based on its reported narrow substrate range, a requirement for calcium for catalytic activity and IP_3being the predominant product from phytate hydrolysis,

Figure 2. Structure of thermostable *Bacillus* sp phytase (Ha *et al*., 2000), a representative model of β-propeller phytase

Oh *et al.* (2004) have proposed that plant alkaline phytases may share a similar catalytic mechanism with BPPhys. Like BPPhys the activity of several plant alkaline phytases, such as lily (*Lilium longiflorum*) pollen (Scott and Loewus, 1986) and a number of legumes (Scott, 1991), is enhanced by calcium. However, none of the plant alkaline phytase genes have been cloned and no sequence data exist to determine if and which of them are BPPhys.

3.3. Cysteine Phosphatase (CP)

It had long been suspected that the presence of certain micro-organisms in the rumen was the reason ruminants could utilize phytic acid and monogastric animals could not. A survey of anaerobic ruminal bacteria has recently revealed phytase activity in one isolate, *Selenomonas ruminantium* (Yanke *et al.*, 1999). The isolation of this micro-organism and the characterization of its unique enzyme have in turn yielded another phytate degrading class of enzymes. Initial characterization established that the enzyme was monomeric, approximately 46 kDa in size, had an optimal pH range of 4.0–5.5, an optimal temperature of 50–55 °C and was inhibited by cations of iron and several other metals. Subsequent studies of this enzyme reveal that it is neither a HAPhy nor a BPPhy (Chu *et al.*, 2004). Its structure and proposed catalytic mechanism suggest it is a member of the cysteine phosphatase (CP) superfamily. Its deduced amino acid sequence contains the active site motif HCXXGXXR(T/S) and other substantial similarities with protein tyrosine phosphatase (PTP), a member of the CP group. The active site forms a loop that functions as a substrate binding pocket unique to PTPs (Fig. 3). The depth of this pocket is important because it appears to determine the substrate specificity (Denu and Dixon, 1998). Consistent with this model *S. ruminantium* phytase, the cysteine phytase (CPhy), has a wider and deeper pocket than PTP and thus is able to accommodate the fully phosphorylated inositol group of phytic acid (Chu *et al.*, 2004). This, plus the presence of a favorable electrostatic environment, allows it to accommodate phytate as a substrate while other members of this group of enzymes lack this capability.

The current model suggests that the initial binding of phytate to the CPhy active site pocket is facilitated by the negatively charged substrate. The hydrolysis of phosphate groups proceeds sequentially with the end product being inositol 2-monophosphate (Chu *et al.*, 2004). The inhibitory effect of iron and other metal cations (Cu^{2+}, Zn^{2+} and Hg^{2+}) has been attributed to their complexing with the substrate, but the stimulatory effect of lead cations remains unexplained (Yanke *et al.*, 1999).

3.4. Purple Acid Phosphatase (PAP)

All members of purple acid phosphatases (PAP) metallophosphoesterases class contain a unique set of seven metal-liganding amino acid residues. These seven metal-liganding residues (D, D, Y, N, H, H, H) are contained in a shared a pattern

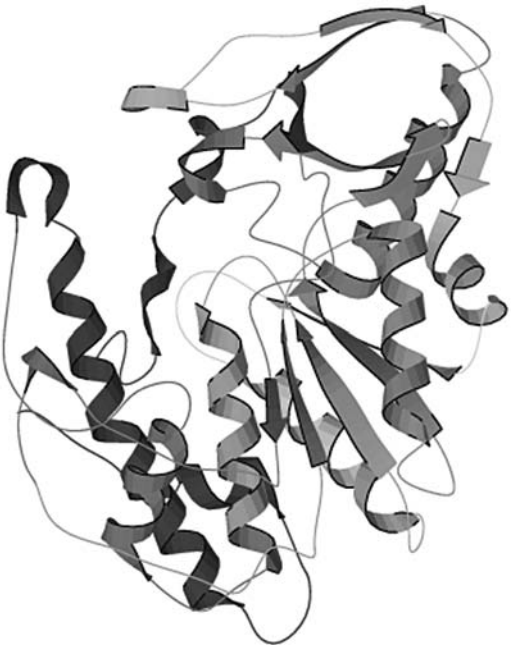


Figure 3. Structure of *Selenomonas ruminantium* phytase (Chu *et al*, 2004), a representative model of cysteine phosphatases

of five common consensus motifs ($DxG/GDx_2Y/GNH(E, D)/Vx_2H/GHxH$) (Schenk *et al*, 2000). This is a large class of phosphtases with known representatives in plants, mammals, fungi and bacteria (Schenk *et al*, 2000; Olczak *et al*, 2003). As in the HAPs and CPs, not all of these enzymes effectively utilized phytate as a substrate. A binuclear metallic center containing two irons is found in animal PAPs, while in plant PAPs the second iron ion is replaced by either a zinc or manganese ion (Olczak *et al*, 2003). The plant PAPs are further divided into two classes: small monomeric proteins (around 35 kDa) and large homodimeric proteins (about 55 kDa). The first binuclear metal-containing hydrolase identified as a phytase was reported in the cotyledons of a germinating soybean (*Glycine max* L. Merr.) seedling (Hegeman and Grabau, 2001). Sequence analysis of GmPhy and other plant PAP phytases (PAPhy) reveal they are similar to the homodimeric plant PAPs.

The adaptation of PAPs in plants to degrade phytate may be a unique case. A fungal PAP, *A. niger* (Apase6) has been isolated and cloned (Ullah and Cummins, 1988; Mullaney *et al*, 1995), but it does not effectively utilize phytate as a substrate. When the active site of A. *niger* Apase 6 and the soybean PAPhys are compared they both contain the conserved active site motif (Mullaney and Ullah, 1998). This indicates the plant PAPhys may have evolved to fill this role as a phytase. However, unlike HAPhys, BPPhys and CPhys, no X-ray crystallography study has

been performed on PAP phytases and no information on the adaptation of the PAPhy active site to phytate as a substrate is available.

Table 1 summarizes the structural division of phytate degrading enzymes based on mechanistic enzymology. This classification scheme furthers the model developed by Oh *et al.* (2004) by incorporating two classes of phytases, PAPhys and CPhys not included in their system. While there are numbers of HAPs, PAPs and CPs that cannot degrade phytic acid effectively, members of each class do share a common enzymatic pathway with other members that do hydrolyze phytate efficiently. How each of these different catalytic mechanisms has evolved to hydrolyze phytate is now being elucidated. This classification system also enables uncharacterized phytases to be assigned based on unique characteristics (calcium requirement, pH profile, end product, *etc*.) associated with each enzyme class. This system also allows for new groups of phytases to be added and existing groups subdivided when desirable.

Table 1. Structural Classes of Phytate Degrading Enzymes

Enzyme Family	Unique Structural Feature	Catalytic mechanism/Adaptation to hydrolyzes Phytate	Examples
Histidine Acid Phosphatase	N-terminal RHGXRXP C-terminal HD consensus motif	N-terminal H forms a phosphohistidine intermediate, C-terminal acts as proton donor/Substrate specificity site residues positively charged	*A. niger* *P. Lycii* *E. coli* *Zea mays* L.
β Propeller Phytase	Six-bladed β propeller shaped molecule	Mechanism consist of an affinity site and a cleavage site. Affinity sites binds phosphate group while other site attacks adjacent phosphate group./Dual site favors IP_6, $IP_{5,\ or}$ IP_4 as substrate.	*Bacillus* sp *X. oryzae*
Cysteine Phosphatase	P loop structure contains HCXXGXXR(T/S) consensus motif	Protein tyrosine phosphatase mechanism cleaves phosphate groups/Deeper active site pocket accommodates phytate	*S. ruminatium*
Purple Acid Phosphatase	Consensus motif: DXG/GDXXY/GNH (E,D) /VXXH/GHXH	Metalloenzymes, phylogenetically linked to large plant PAP/unknown	*Glycine max* *M. truncatula*

4. APPLICATIONS OF PHYTASE

Since the first commercial phytase product Natuphos® was launched in 1991, the market volume has reached ca. 150 million euros and will likely expand with new applications. The main application is still as a feed supplement to improve P bioavailability in plant feed-stuffs via the enzyme-mediated hydrolysis of phytate. Most importantly, the improved utilization of the phosphate deposits in the feed results in a substantial reduction in the phosphate content in animal manure and hence decreases of phosphate load on the environment in areas of intensive animal agriculture. High dietary P bioavailability reduces the need for supplemental inorganic phosphorus such as mono- and dicalcium-phosphate (MCP, DCP). Because of the strong economic growth in China and India along with the oil price hike, the supply and cost of MCP and DCP has become a practical issue. Furthermore, inorganic phosphate is non-renewable resource, and it has been estimated that the easily-accessible phosphate on earth will be depleted in 50 years. Thus, phytase is an effective tool for natural resource management of phosphorus on a global scale.

The ban of dietary supplementation of meat and bone meal, as a cheap source of feed phosphorus, in Europe to prevent possible cross-species transfer of diseases such as BSE, has led to a profound change in the feed P management. This has given phytase a new socio-economic impact as a cost effective alternative to ensure animals to obtain adequate available P from the plant-based diets. Being the major storage form of P in seeds, plant phytate was produced in 2000 at a global yield > 51 million metric tons. This amount accounts for approximately 65% of the elemental P sold world wide as fertilizers (Lott *et al*, 2000). Apparently, phytase can turn the plant phytate into a very valuable resource of P by improving its bioavailability for animal nutrition. Denmark and the Netherlands have imposed regulations to promote the use of microbial phytases.

4.1. As a Feed Supplement

Supplemental dietary phytases have been shown to effectively improve phytate-P utilization to simple-stomached animals under various dietary conditions. In a maize based diet with little intrinsic phytase activity, the improvement derived by the supplemental phytase is generally greater than that in a barley/wheat-based diet with a significant phytase activity. In pigs, supplemental microbial phytase at 750 units/kg increased P bioavailability in maize from 18 to 56%, in wheat from 62 to 74%, and in triticale from 52 to 67% (Dungelhoef *et al*, 1994). Based on large bodies of literature, Ravidran (1995) and Maenz (2001) concluded that supplementation with microbial phytase to diets for poultry and pigs enhanced phytate-P utilization by 20 to 45%. Consequently, fecal phosphorus excretion is reduced by 42% in weanling pigs (Lei *et al*, 1993a). The P equivalency of microbial phytase for 1 g of non-phytate P in broilers is equivalent to 650 to 750 FTU phytase $\times$ kg^{1} diet (Kornegay *et al*., 1996; Schoner *et al*, 1991). In pigs, 1 g of inorganic P is equivalent to 500 FTU $\times$ kg^{-1} microbial phytase (Yi *et al*, 1996).

Phytic acid is a strong anti-nutritional factor because of the ability of phytic acid and the lower myo-inositol phosphates to form complexes with divalent metals. Thus, supplemental phytase has been shown to improve bioavailabilities of minerals (Lei *et al*., 1993a, b). Jondreville *et al*. (2005) reported that 500 units of Natuphos® replaced 30 mg of Zn as sulphate given in a maize-soya-bean meal based diet. However, increasing Ca levels from 0.4 to 0.8% in low P maize-soybean meal diets significantly reduced the efficacy of microbial phytase in weanling pigs (Lei *et al*., 1994). Effects of supplemental microbial phytases on digestibility of starch, protein, and amino acids have been reported, but inconsistent.

4.2. Potential in Human Nutrition

As mentioned above, ingesting high levels of dietary phytate severely impedes the absorption of important trace elements such as iron and zinc in digestive tracts. Partially due to this anti-nutritional effect of phytate, approximately two to three billion people, primarily in the developing world, afflict deficiencies of these nutrients. There are two ways to reduce dietary phytate intake and its negative effects. One is to develop low-phytate crops via impairing the phytic acid biosynthesis by disruption of the inositol polyphosphate kinases or other mutations (Brinch-Pedersen *et al*., 2002; Stevenson-Paulik *et al*., 2005). Although this approach has led to success for the primary goal, the low-phytate maize and soybean have shown reductions in yield and seed germination (Pilu *et al*., 2003; Oltmans *et al*., 2005). This undesirable changes may not be completely surprising as recent evidence shows that phytate is required as a cofactor for RNA editing (Macbeth *et al*., 2005), in addition to its previously recognized roles in storing P and energy for germination.

Supplementing phytogenic or extrinsic phytases into human foods is another, perhaps more effective, way to reduce the negative effect of phytate on mineral utilization. Both sources of phytases were effective in reducing bread phytate and improving iron availability (Porres *et al*., 2001). The phytase-mediated dephytinization of infant formulas, infant cereals or complementary foods has been studied, and the effectiveness in improving trace element nutrition is greater if the protein supply is from legume seeds instead of milk (Hurrell, 2004). Treatments with phytase, or in combination with other processing, have been assayed in cereal and legumes (Maklinder *et al*., 1995; Porres *et al*., 2005) for developing food ingredients with high nutritional and functional value.

There are issues to be clarified before implementing a phytase strategy in human nutrition. Although activation of the intrinsic phytase, *i.e.* during baking of whole wheat bread can decrease the phytate content by 50 to 60%, some research has indicated that the reduction must reach $> 77\%$ to show impacts on iron and zinc absorption in humans (for review see Sandberg and Andlid, 2002). Meanwhile, certain levels of dietary inositol phosphates may be health-beneficial with possible functions in antioxidation, anti-tumorigenesis, reducing serum lipids and cholesterol levels, preventing renal calculi via mineral-binding, and in diabetes

(Burgess and Gad, 2002; Jenab and Thompson, 2002). It will be a challenge to minimize the negative effect of phytate on iron and zinc nutrition without losing its potential health benefits.

4.3. Novel Industrial Uses

In consideration of the potential health values of certain inositol phosphates, phytase may be used in cost-effective bioreactors for large scale production of these compounds. Successful attempts to immobilize phytase on a variety of matrices have been made. Similarly, Quan *et al.* (2003) have attempted the immobilization of a phytase-producing *Candida crusei* cells on calcium alginate gel-beads for the preparation of specific *myo*-inositol phosphates. Variation in the composition of end products resulted from a change in the flow rate of phytic acid solution (5mM) through the bioreactor.

Fujita *et al.* (2001) have tested a mutant strain of *A. oryzae* with high phytase activity for sake brewing. When compared to the wild-type strain, alcohol fermentation was promoted in the high phytase producing strain with a subsequent increase in the yield of alcohol production. Haros *et al.* (2001) have used exogenous microbial phytase as a novel bread making additive to improve several baking and physical parameters like proofing time (24% reduction) width/height ratio of bread slices (5% reduction), specific volume (21% increase), and crumb firmness (28.3% reduction). A novel industrial use of phytase has been proposed for efficient separation of soybean β-conglycinin and glycinin by Saito *et al.* (2001). Although commercial use remains to be tested, the potential role of thermophilic phytases as powerful additives in the pulp and paper industry has been suggested (Madhavan *et al.*, 2004). Phytase could act synergistically with xylanase in crude multi-enzyme preparation from xylanase-producing micro-organisms like *Streptomyces cupidosporus* (Maheswari, 2000) that are used for the treatment of pulp.

Enzyme immobilization has been employed by Vieira and Nogeira (2004) for the development of a flow injection spectrophotometric procedure to determine the amount of ortho-phosphate, phytate and total phosphorus in cereal samples, and by Mak *et al.* (2004) for the development of novel biosensors to measure phytic acid and phytase activity.

5. CURRENT RESEARCH INTERESTS

5.1. Improving Heat Stability

Most phytases from plants and micro-organisms start to lose activity around 55–60 °C (Phillippy, 2002). For example, the T_m of *A. niger phy*A phytase is 63.3 °C and the denaturation is associated with an irreversible conformational change with loss of 70% to 80% of the activity. Apparently, the limited thermotolerance is the major constraint for the application of phytase in both feed and food industry because most animal feed are pelleted at 80–90 °C in order to eliminate

Salmonella infections, and human staples are processed by boiling or baking. Therefore, intensive research and development have been directed toward screening of thermophiplilic and hyperthermophilic organisms for thermotolerant enzymes, mutagenesis of mesophilic enzymes to increase their thermostability, and design of formulations of chemical coating for protecting the enzymes from heat-denaturation. Initial success has resulted from rational design to generate a consensus phytase with a T_m of 89.3 °C (Lehmann *et al*., 2000a). The initial enzyme originating from *A. fumigatus* is not a genuine heat stable enzyme, as it posses a low T_m (62.5 °C) but can refold into a fully active conformation after cooling (Wyss *et al*., 1998).

Dozens of exogenous agents have been reported to stabilize phytase during pelleting (see Phillippy, 2002 for review). However, endogenous seed components also have a protective effect on phytase activity. Skoglund *et al*. (1997) found that the intrinsic phytase in a barley-pea-rapeseed pig diet was unaffected after pelleting at 81 °C. When transgenic rice expressing the *A. fumigatus* phytase was boiled for 20 min, only 8% of the initial phytase activity was retained (Lucca *et al*., 2001). This residual activity was much lower than that (59%) of the commercial preparation of the fungal enzyme boiled for the same period of time, indicating that the *in planta* expression may interfere with refolding of the enzyme or provide an environment unfavourable to refolding. Expression and accumulation of the consensus phytase 10-thermo-[3]-Q50T-K91A and the *A. fumigatus* phytase in the protein storage vacuoles of transgenic wheat has recently revealed that with this inter-cellular deposition, heat stability based on high unfolding temperature of the consensus is superior to high refolding capacity of the *A. fumigatus* enzyme (Brinch-Pedersen *et al*., 2006).

5.2. Shifting pH-Activity Profiles

Since many of the known phytases have pH optima that are not within the pH range of the stomach (the main site of phytate-hydrolysis), approaches have been taken to improve pH-activity profiles. Those include modification of ionizable groups directly involved in substrate specificity or catalysis, replacement of amino acid residues in direct contact with residues located in the active or substrate specificity site by means of hydrogen bonds or salt bridges, or alteration of distant charge interactions by modification of the surface charge of the enzyme. Using these strategies, Tomschy *et al*. (2002) improved the activity of *A. fumigatus* phytase and a consensus phytase at low pH.

To improve the catalytic properties of the thermostable consensus-1 phytase, Lehmann *et al*. (2000b) replaced all the divergent amino acid residues present in the active site of the consensus phytase by those of *A. niger NRRL 3135* phytase. The new phytase termed consensus-7 phytase featured a major shift in the catalytic properties that were similar to those of *A. niger NRRL 3135* phytase, thus demonstrating the feasibility of rational transfer of favorable catalytic properties. However, the active site residues transfer caused a decrease in the unfolding temperature of consensus-7 phytase compared with consensus-1 phytase.

5.3. Enhancing Acid and Protease Resistance

The low pH in the stomach (2–5) of pigs and poultry, in the crop of poultry (4–5) and the neutral pH in the small intestine (6.5–7.5) provide challenges for the stability of feed phytases (Simon and Igbasan, 2002; Phillippy, 2002). Konietzny (2002) concluded that most microbial phytases are more pH stable than plant phytases as stability of most plant phytases was reduced significantly at $4 < \text{pH} > 7.5$, whereas most microbial enzymes are relatively stable at $3 < \text{pH} > 8.0$.

An effective phytase needs to be resistant to proteolysis in the animal digestive tracts. Major differences in the resistance to pancreatic and pepsin proteases are present among various phytases. Phillippy (1999) reported that intrinsic wheat phytase is more susceptible to inactivation by pancreatin and pepsin than *A. niger* phytase. In contrast, the *A. niger* phytase is less resistant to degradation by pepsin than recombinant *E. coli* phytase (r-AppA) (Rodriguez *et al*, 1999a). In another comparison of *Bacillus subtilis* and *E. coli* with four recombinant fungal phytases, the fungal phytases were most susceptible to inactivation by pancreatin and pepsin, *Bacillus subtilis* was stable to pancreatin whereas *E. coli* was stable to both pancreatin and pepsin (Igbasan *et al*, 2000).

Commercial phytases also need to resist degradation during production and storage. When the phytases from *A. fumigatus* and *Emericella nidulans* were expressed in *A. niger*, they were cleaved by proteases present in the culture supernatant (Wyss *et al*, 1999b). The cleavage had no effect on the activity of *E. nidulans* phytase, but significantly reduced the activity of *A. fumigatus* phytase activity. Site-directed mutagenesis at the protease sensitive sites of the *A. fumigatus* and *E. nidulans* phytases yielded mutants with significant reduced susceptibility to proteolytic degradation.

5.4. Searching for Efficient Production Systems

Phytase production has been attempted in several fungal species by either submerged or solid state fermentation. Fungal species from the *Aspergillus* genera are widely employed for phytase production (Wyss *et al*, 1999a, b; Martin *et al*, 2003), although other mesophilic fungi like *Mucor racemosus* and *Rhizopus oligosporus* (Sabu *et al*, 2002; Bogar *et al*, 2003), or the thermophilic fungi *T. aurantiacus* (Madhavan *et al*, 2004) have also been employed.

Methylotrophic yeast such as *P. pastoris* or *Hansenula polymorpha* exhibit great potential for producing high levels of *A. niger, E. coli*, and *A. fumigatus* phytases (Mayer *et al*, 1999; Rodriguez *et al*, 1999a, b, 2000a, b). The phytase production can be greatly enhanced by optimizing culture conditions, restriction of oxygen supply during passage, stabilization and screening process, and changes in codon usage of the phytase gene and modification of signal peptide (Mayer *et al*, 1999; Stöckmann *et al*, 2003).

To circumvent the inconvenience of intracellular expression, Miksch *et al* (2002) have tested extracellular expression of *E. coli* phytase in bacteria using a secretion

system based on the controlled expression of the *kil* gene. Arndt *et al.* (2005) have further developed a controller based on glucose detection from the culture media and implementation of a kalman filter in order to maximize extracellular phytase expression by *E. coli* and minimize acetate production that inhibits cell growth. Kerovuo *et al.* (2000) have developed a novel *Bacillus* expression system for efficient extracellular phytase production, whereas Gerlach *et al.* (2004) have attempted optimal expression of *E. coli* phytase in *Bacillus subtilis* bearing induced *Tat*-dependent transport system components (*TatAd/TatCd*) using the PhoD-specific export signals. Introduction of the *appA* signal peptide cleavage site into the fusion protein resulted in efficient processing of the recombinant enzyme, and expression was further improved by the use of a protease deficient *Bacillus subtilis wprA* strain.

In searching for alternative efficient bacterial expression systems, Lan *et al.* (2002) have studied the effect of different components of culture media and culture conditions on phytase production by the rumen bacteria *Mitsuokella jalaludinii* in fed-batch fermentation. Dharmsthiti *et al.* (2005) have successfully attempted *E. coli* appA phytase production in *Pseudomonas putida* based on the effective secretory system, less limitations in codon usage, and capability of growth using a wide variety of substrates. Effective extracellular *E. coli AppA* phytase expression has also been reported by Stahl *et al.* (2003) in *Streptomyces lividans*. Recently, Yin *et al.* (2005) have described a novel phytase expression system based on silkworm larvae with a yield of 7,710 units per ml hemolymph.

5.5. Developing Phytase Transgenic Plants and Animals

Phytase transgenic pigs, named EnviroPig, have been generated by overexpressing the *E. coli* AppA phytase in their saliva, and have shown reductions in fecal P output by up to 75% (Golovan *et al.*, 2001). The transgene phytase strategy has also been implemented in several crops (see Brinch-Pedersen *et al.*, 2002 for review). The phytase genes of choice in these plant studies include those from *A. niger, S. ruminantium*, *E. coli* and *A. fumigatus* or synthetic ones.

While the first studies were based on constitutive expression driven by the cauliflower mosaic virus (CaMV) 35S and maize ubiquitin promoters, subsequent studies have largely used seed specific expression promoters such as the wheat HMW glutenine 1DX5, the oilseed rape cruciferin and the rice glutelin promoters. Germination-specific expression directed by the specific α-amylase α*amy*8 promoter, was used in rice. To ensure guidance into the ER for glycosylation and export to the appoplast, signal peptides from tobacco extensin, barley α-amylase, *Brassica napus* cruceferin, barley β-glucanase, rice α-amylase and soybean VSPβ were tested. Recent studies in wheat have unraveled unequivocally that the heterologous phytase is deposited in the vacuole albeit that the transformation constructs were designed for secretion to the appoplast (Brinch-Pedersen *et al.*, 2006). A microarray study of transgenic wheat has shown no unintended side effects on the wheat transcriptome by the expressing of an *A. fumigatus* phytase (Gregersen *et al.*, 2005).

Phytase transgenic seeds of soybean (Denbow *et al*., 1998), oilseed rape (Zhang *et al*., 2000a, b) and tobacco (Pen *et al*., 1993) have been tested in feeding trials with broilers, pigs, and rats. Without any unfavorable feeding effects, these phytase transgenic seeds have improved P utilization and reduced manure P excretion, comparably with supplemental microbial phytases.

ACKNOWLEDGEMENTS

The phytase research in Lei's Laboratory was funded in part by Cornell Biotechnology Program. Jesus Porres is working under a research contract from Junta de Andalucia, Spain and Project AGL2002-02905 ALI.

REFERENCES

Arndt, M., Kleist, S., Miksch, G., Friehs, K., Flaschel, E., Trierweiler, J. and Hitzmann, B. (2005). A feedforward-feedback substrate controller based on a kalman filter for a fed-batch cultivation of *Escherichia coli* producing phytase. Comput. Chem. Eng. *29*, 1113–1120.

Augspurger, N.R., Webel, D.M., Lei, X.G. and Baker, D.H. (2003). Efficacy of an *E. coli* phytase expressed in yeast for releasing phytate-bound phosphorus in young chicks and pigs. J. Anim. Sci. *81*, 474–483.

Berka, R.M., Rey, M.W., Brown, K.M., Byun, T. and Klotz, A.V. (1998). Molecular characterization and expression of a phytase gene from the thermophilic fungus *Thermomyces lanuginosus*. Appl. Environ. Microbiol. *64*, 4423–4427.

Bitar, K. and Reinhold, J.G. (1972). Phytase and alkaline phosphatase activities in intestinal mucosae of rat, chicken, calf, and man. Biochim. Biophys. Acta. *268*, 442–452.

Bogar, B., Szakacs, G., Linden, J.C., Pandey, A. and Tengerdy, R.P. (2003). Optimization of phytase production by solid substrate fermentation. J. Ind. Microbiol. Biotechnol. *30*, 183–189.

Brinch-Pedersen, H., Hatzack, F., Stoger, E., Arcalis, E. and Holm, P.B. (2006). Heat stable phytases in transgenic wheat (*Triticum aestivum* L.): deposition pattern, thermostability and phytate hydrolysis. J. Agric. Food Chem. In Press.

Brinch-Pedersen, H., Sorensen, L.D. and Holm, P.B. (2002). Engineering crop plants: getting a handle on phosphate. Trends. Plant. Sci. *7*, 118–125.

Burgess, J.R. and Gao, G. (2002). The antioxidant effects of inositol phosphates. In Food Phytates. Eds. Reddy, A.R. and Sathe. S.K. CRC Press, Bocca Raton, FL, USA. pp. 189–197.

Chadha, B.S., Harmeet, G., Mandeep, M., Saini, H.S. and Singh, N. (2004). Phytase production by the thermophilic fungus *Rhizomucos pusillus*. World. J. Microbiol. Biotechnol. *20*, 105–109.

Cheng, C. and Lim, B.L. (2005). Beta-propeller phytases in the aquatic environment: characterization of a novel phytase from *Shewanella oneidensis* MR-1. In Inositol Phosphates in the Soil-Plant-Animal System: Linking Agriculture and Environment. Proceedings of the Bouyoucos Conference to Address the Biogeochemical Interaction of Inositol Phosphates in the Environment; Turner, B.L., Richardson, A.E., Mullaney, E.J., Eds.; Sun Valley, Idaho, USA. pp. 55–56.

Cho, J.S., Lee, C.W., Kang, S.H., Lee, J.C., Bok, J.D., Moon, Y.S., Lee, H.G., Kim, S.C. and Choi, Y.J. (2003). Purification and characterization of a phytase from *Pseudomonas syringae* MOK1. Curr. Microbiol. *47*, 290–294.

Chu, H.M., Guo, R.T., Lin, T.W., Chou, C.C., Shr, H.L., Lai, H.L., Tang, T.Y., Cheng, K.J., Selinger, B.L. and Wang, A.H.J. (2004) Structures of *Selenomonas ruminantium* phytase in complex with persulfated phytate: DSP phytase fold and mechanism for sequential substrate hydrolysis. Structure *12*, 2015–2024

Denbow, D.M., Grabau, E.A., Lacy, G.H., Kornegay, E.T., Russell, D.R. and Umbeck. P.F. (1998). Soybeans transformed with a fungal phytase gene improve phosphorus availability for broilers. Poult. Sci. *77*, 878–881.

Denu, J.M. and Dixon, J.E. (1998) Protein tyrosine phosphatase: mechanism of catalysis and regulation. Cur. Opinion Chem. Biol. *2*, 633–641.

Dharmsthiti, S., Chalermpornpaisarn, S., Kiatiyajarn, M., Chanpokapaiboon, A., Klongsithidej, Y. and Techawiparut, J. (2005). Phytase production from *Pseudomonas putida* harboring *Escherichia coli appA*. Process. Biochem. *40*, 789–793.

Dungelhoef, M., Rodehutscord, M., Spiekers, H. and Pfeffer, E. (1994). Effects of supplemental microbial phytase on availability of phosphorus contained in maize, wheat and triticale to pigs. Anim. Feed Sci. Technol. *49*, 1–10.

Ellestad, L.E., Angel, R. and Soares Jr, J.H. (2002). Intestinal phytase II: A comparison of activity and *in vivo* phytate hydrolysis in three teleost species with differing digestive strategies. Fish Physiol. Biochem. *26*, 259–273.

Fredlund, K., Bergman, E.L., Rossander-Hulthen, L., Isaksson, M., Almgren, A. and Sandberg, A.S. (2003). Hydrothermal treatment and malting of barley improve zinc absorption but not calcium absorption in humans. Eur. J. Clin. Nutr. *57*, 1507–1513.

Freund, W.D., Mayr, G.W., Tietz, C. and Schultz, J.E. (1992). Metabolism of inositol phosphates in the protozoan *Paramecium*. Characterization of a novel inositol-hexakisphosphate-dephosphorylating enzyme. Eur. J. Biochem. *207*, 359–367.

Fujita, J., Fukuda, H., Yamance, Y., Kizaki, Y., Shigeta, S., Ono, K., Suzuki, O. and Wakabayashi, S. (2001). Critical importance of phytase for yeast growth and alcohol fermentation in Japanese *sake* brewing. Biotechnol. Lett. *23*, 867–871.

Gerlach, R., Pop, O. and Müller, J.P. (2004). Tat dependent export of *E. coli* phytase AppA by using the PhoD-specific transport system of *bacillus subtilis*. J. Basic. Microbiol. *5*, 351–359.

Golovan, S.P., Meidinger, R.G., Ajakaiye, A., Cottrill, M., Wiederkehr, M.Z., Barney, D.J., Plante, C., Pollard, J.W., Fan, M. Z,, Hayes, M.A., Laursen, J., Hjorth, J.P., Hacker, R.R., Phillips, J.R. and Forsberg, C.W. (2001). Pigs expressing salivary phytase produce low-phosphorus manure. Nature Biotechnol. *19*, 741–745.

Gregersen, P.L., Brinch-Pedersen, H. and Holm, P.B. (2005). A microarray based, comparative analysis of the gene expression profiles during grain development in transgenic and wild type wheat. Transgenic Res. *14*, 887–905.

Greiner, R., Konietzny, U. and Jany, K.D. (1993). Purification and characterization of two phytases from *Escherichia coli*. Arch. Biochem. Biophys. *303*, 107–113.

Greiner, R., Larsson Alminger, M., Carlsson, N.G., Muzquiz, M., Burbano, C., Cuadrado, C., Pedrosa, M. and Goyoaga, C. (2002). Pathway of dephosphorylation of *myo*-inositol hexakisphosphate by phytases of legume seeds. J. Agric. Food Chem. *50*, 6865–6870.

Ha, N.C., Oh, B.C., Shin, S., Kim, H.J., Oh, T.K., Kim, Y.O., Choi, K.Y. and Oh, B.H. (2000). Crystal structures of a novel, thermostable phytase in partially and fully calcium-loaded states. Nature. Struct. Biol. *7*, 147–153.

Han, Y.M., Wilson, D.B. and Lei, X.G. (1999). Expression of an *Aspergillus niger* phytase gene (*phyA*) in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. *65*, 1915–1918.

Han, Y.M., Yang, F., Zhou, A.G., Miller, E.R., Ku, P.K., Hogberg, M.G. and Lei, X.G. (1997). Supplemental phytases of microbial and cereal sources improve dietary phytate phosphorus utilization by pigs from weaning through finishing. J. Anim. Sci. *75*, 1017–1025.

Haros, M., Bielecka, M. and Sanz, Y. (2005). Phytase activity as a novel metabolic feature in *Bifidobacterium*. FEMS. Microbiol. Lett. *247*, 231–239.

Haros, M., Rosell, C.M. and Benedito, C. (2001). Use of fungal phytase to improve breadmaking performance of whole wheat bread. J. Agric. Food Chem. *49*, 5450–5454.

Hegeman, C.E. and Grabau, E.A. (2001). A novel phytase with sequence similarity to purple acid phosphatases is expressed in cotyledons of germinating soybean seedlings. Plant. Physiol. *126*,1598–1608.

Hurrell, R.F. (2004). Phytic acid degradation as a means of improving iron absorption. Int. J. Vitam. Nutr. Res. *74*, 445–452.

Igbasan, F.A., Manner, K., Miksch, G., Borriss, R., Farouk, A. and Simon, O. (2000). Comparative studies on the *in vitro* properties of phytases from various microbial origins. Arch. Anim. Nutr. *53*, 353–373.

Jenab, M. and Thompson, L.U. (2002). Role of phytic acid in cancer and other diseases. In Food Phytates. Eds. Reddy, A.R. and Sathe, S.K. CRC Press, Boca Raton, FL, USA. pp. 225–248.

Jog, S.P., Garchow, B.G., Mehta, B.D. and Murthy, P.P.N. (2005). Alkaline phytase from lily pollen: Investigation of biochemical properties. Arch. Biochem. Biophys. *440*, 133–140.

Jondreville, C., Hayler, R. and Feuerstein, D. (2005). Replacement of zinc sulphate by microbial phytase for piglets given a maize-soya-bean meal diet. J. Anim. Sci. *81*, 77–83.

Kerovuo, J., Lauraeus, M., Nurminem, P., Kalkkinen, N. and Apajalahti, J. (1998). Isolation, characterization, molecular gene cloning, and sequencing of a novel phytase from *Bacillus subtilis*. Appl. Environ. Microbiol. *64*, 2079–2085.

Kerovuo, J., Von Weymarn, N., Povelainen, M., Auer, S. and Miasnikov, A. (2000). A new efficient expression system for *Bacillus* and its application to production of recombinant phytase. Biotechnol. Lett. *22*, 1311–1317.

Kim, H., Kim, Y., Lee, J., Kim, K. and Kim, Y. (2003). Isolation and characterization of a phytase with improved properties from *Citrobacter braakii*. Biotechnol. Lett. *25*, 1231–1234.

Kim, Y.O., Kim, H.K., Bae, K.S., Yu, J.H. and Oh, T.K. (1998) Purification and properties of a thermostable phytase from *Bacillus* sp. DS11. Enzyme Microb. Technol. *22*, 2–7.

Konietzny, U. and Greiner, R. (2002). Molecular and catalytic properties of phytate-degrading enzymes (phytases). Int. J. Food Sci. Tech. *37*, 791–812.

Kornegay, E.T., Denbow, D.M., Yi, Z. and Ravindran, V. (1996). Response of broilers to graded levels of microbial phytase added to maize-soyabean-meal-based diets containing three levels of non-phytate phosphorus. Br. J. Nutr. *75*, 839–852.

Kostrewa, D., Grueninger-Leitch, F., D'Arcy, A., Broger, C., Mitchell, D. and van Loon, A.P.G.M. (1997). Crystal structure of phytase from *Aspergillus ficuum* at 2.5 A resolution. Nature Struc. Biol. *4*, 185–190.

Kostrewa, D., Wyss, M., D'Arcy, A. and van Loon, A.P.G.M. (1999) Crystal structure of *Aspergillus niger* pH 2.5 acid phosphatase at 2.4 A resolution. J. Mol. Biol. *288*, 965–974.

Lan, G.Q., Abdullah, N., Jalaudin, S. and Ho, Y.W. (2002). Culture conditions influencing phytase production of *Mitsuokella jalaudinii*, a new bacterial species from the rumen of cattle. J. Appl. Microbiol. *93*, 668–674.

Lassen, S.F., Breinholt, J., Østergaard, P.R., Brugger, R., Bischoff, A., Wyss, M. and Fuglsang, C.C. (2001). Expression, gene cloning, and characterization of five novel phytases from four basidiomycete fungi: *Peniophora lycii*, *Agrocybe pediades*, a *Ceriporia* sp. and *Trametes pubescens*. Appl. Environ. Microbiol. *67*, 4701–4707.

Lehmann, M., Kostrewa, D., Wyss, M., Brugger, R., D'Arcy, A., Pasamontes, L. and van Loon, A.P.G.M. (2000a). From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase. Prot. Eng. *13*, 49–57.

Lehmann, M., Lopez-Ulibarri, R., Loch, C., Viarouge, C., Wyss, M. and van Loon, A.P.G.M. (2000b). Exchanging the active site between phytases for altering the functional properties of the enzyme. Prot. Sci. *9*, 1866–1872.

Lei, X.G., Ku, P.K., Miller, E.R., Yokoyama, M.T. and Ullrey, D.E. (1993a). Supplementing corn-soybean meal diets with microbial phytase maximizes phytate phosphorus utilization by weanling pigs. J. Anim. Sci. *71*, 3368–3375.

Lei, X.G., Ku, P.K., Miller, E.R., Ullrey, D.E. and Yokoyama, M.T. (1993b). Supplemental microbial phytase improves bioavailability of dietary zinc to weanling pigs. J. Nutr. *123*, 1117–1123.

Lei, X.G., Ku, P.K., Miller, E.R., Yokoyama, M.T. and Ullrey, D.E. (1994). Calcium level affects the efficacy of supplemental microbial phytase in corn-soybean meal diets of weanling pigs. J. Anim. Sci. *72*, 139–143.

Lim, D., Golovan, S., Forsberg, C.W. and Jia, Z. (2000) Crystal structure of *Escherichia coli* phytase and its complex with phytate. Nature Struct. Biol. *7*, 108–113.

Liu, Q., Huang, Q., Lei, X.G. and Hao, Q. (2004) Crystallographic snapshots of *Aspergillus fumigatus* phytase, revealing its enzymatic dynamics. Structure *12*, 1578–1583.

Lott, J.N.A., Ockenden, I., Raboy, V. and Batten, G.D. (2000). Phytic acid and phosphorus in crop seeds and fruits: a global estimate. Seed Sci. Res. *10*, 11–33.

Lucca, P., Hurrell, R. and Potrykus, I. (2001). Genetic engineering approaches to improve the bioavailability and the level of iron in rice grains. Theor. Appl. Genet. *102*, 392–397.

Macbeth, M.R., Schubert, H.L., VanDemark, A.P., Lingam, A.T., Hill, C.P. and Bass, B.L. (2005). Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. Science *309*, 1534–1539.

Madhavan Nampoothiri, K., Jino Tomes, G., Roopesh, K., Szakacs, G., Nagy, V., Soccol, C.R. and Pandey, A. (2004). Thermostable phytase production by *Thermoascus aurantiacus* in submerged fermentation. Appl. Biochem. Biotechnol. *118*, 205–214.

Maenz, D.D. (2001). Enzymatic characteristics of phytases as they relate to their use in animal feed. In Enzymes in farm animal nutrition. Eds. Bedford, M.R. and Partridge, G.G. CABI Publishing, Wallington, UK. pp. 61–84.

Maenz, D.D. and Classen, H.L. (1998). Phytase activity in the small intestinal brush border membrane of the chicken. Poult. Sci. *77*, 557–563.

Maheswari, M.U. and Chandra, T.S. (2000). Production and potential application of a xylanase from a new strain of *Streptomyces cupidosporus*. World. J. Microbiol. Biotechnol. *16*, 257–263.

Mak, W.C., Ng, Y.M., Chan, C., Kwong, W.K. and Renneberg, R. (2004). Novel biosensors for quantitative phytic acid and phytase measurement. Biosens. Bioelect. *19*, 1029–1035.

Maklinder, I.M., Larsson, M., Fredlund, K. and Sandberg, A.S. (1995). Degradation of phytate by using varied sources of phytases in an oat-based nutrient solution fermented by *Lactobacillus plantarum* strain 299 V. Food Microbiol. *12*, 487–495.

Martin, J.A., Murphy, R.A. and Power, R.F.G. (2003). Cloning and expression of fungal phytases in genetically modified strains of *Aspergillus awamori*. J. Ind. Microbiol. Biotechnol. *30*, 568–576.

Mayer, A.F., Hellmuth, K., Schlieker, H., Lopez-Ulibarri, R., Oertel, S., Dahlems, U., Strasser, A.W.M. and van Loon, A.P.G.M. (1999). An expression system matures: a highly efficient and cost-effective process for phytase production by recombinant strains of *Hansenula polymorpha*. Biotechnol. Bioeng. *63*, 373–381.

Miksch, G., Kleist, S., Friehs, K. and Flaschel, E. (2002). Overexpression of the phytase from *Escherichia coli* and its extracellular production in bioreactors. Appl. Microbiol. Biotechnol. *59*, 685–694.

Mitchell, D.B., Vogel, K., Weimann, B.J., Pasamontes, L. and van Loon, A.P.G.M. (1997). The phytase subfamily of histidine acid phosphatases: isolation of genes for two novel phytases from the fungi *Aspergillus terreus* and *Myceliophthora thermophila*. Microbiol. *143*, 245–252.

Mullaney, E.J., Daly, C.B., Ehrlich, K.C. and Ullah, A.H.J. (1995) The *Aspergillus niger (ficuum) aphA* gene encodes a pH 6.0-optimum acid phosphatase. Gene *162*, 117–121.

Mullaney, E.J., Daly, C.B., Kim, T., Porres, J.M., Lei, X.G., Sethumadhavan, K. and Ullah, A.H.J. (2002) Site-directed mutagenesis of *Aspergillus niger* NRRL 3135 phytase at residue 300 to enhance catalysis at pH 4.0. Biochem. Biophys. Res. Commun. *297*, 1016–1020.

Mullaney, E.J., Daly, C.B. and Ullah, A.H.J. (2000) Advances in phytase research. Adv. Appl. Microbiol. *47*, 157–199.

Mullaney, E.J. and Ullah, A.H.J. (1998) Conservation of the active site motif in *Aspergillus niger (ficuum)* pH 6.0 optimum acid phosphatase and kidney bean purple acid phosphatase. Biochem. Biophys. Res. Commun. *243*, 471–473.

Mullaney, E.J. and Ullah, A.H.J. (2003) The term phytase comprises several different classes of enzymes. Biochem. Biophys. Res. Commun. *312*, 179–184.

Mullaney, E.J. and Ullah, A.H.J. (2005) Conservation of cysteine residues in fungal histidine acid phytases. Biochem. Biophys. Res. Commun *328*, 404–408.

Nakamura, Y., Fukuhara, H. and Sano, K. (2000). Secreted phytase activities of yeasts. Biosci. Biotechnol. Biochem. *64*, 841–844.

Oh, B.C., Chang, B.S., Park, K.H., Ha, N.C., Kim, H.K., Oh, B.H. and Oh, T.K. (2001) Calcium-dependent catalytic activity of a novel phytase from *Bacillus amyloloquefaciens* DS11. Biochem. *40*, 9669–9676.

Oh, B.C., Choi, W.C., Park, S., Kim, Y.O. and Oh, T.K. (2004). Biochemical properties and substrate specificities of alkaline and histidine acid phosphatase. Appl. Microb. Biotechnol. *63*, 362–372.

Olczak, M., Morawiecha, B. and Watorek, W. (2003) Plant purple acid phosphatases- genes, structures and biological function. Acta Biochimica Polonica. *50*, 1245–1256.

Oltmans, S.E., Fehr, W.R., Welke, G.A., Raboy, V. and Peterson, K.L. (2005). Agronomic and seed traits of soybean lines with low-phytate phosphorus. Crop. Sci. *45*, 593–598.

Pasamontes, L., Haiker, M., Henriquez-Huecas, M., Mitchell, D.B. and van Loon, A.P.G.M. (1997b). Cloning of the phytase from *Emericella nidulans* and the thermophilic fungus *Talaromyces thermophilus*. Biochim. Biophys. Acta. *1353*, 217–223.

Pasamontes, L., Haiker, M., Wyss, M., Tessier, M. and van Loon, A.P.G.M. (1997a). Gene cloning and characterization of a heat-stable phytase from the fungus *Aspergillus fumigatus*. Appl. Environ. Microbiol. *63*, 1696–1700.

Pen, J., Verwoerd, T.C., Vanparidon, P.A., Beudeker, R.F., Vandenelzen, P.J.M., Geerse, K., Vanderklis, J.D., Versteegh, H.A.J., Vanooyen, A.J.J. and Hoekema, A. (1993). Phytase-containing transgenic seeds as a novel feed additive for improved phosphorus utilization. Bio-Technology. *11*, 811–814.

Phillippy, B.Q. (1999). Susceptibility of wheat and *Aspergillus niger* phytases to inactivation by gastrointestinal enzymes. J. Agric. Food Chem. *47*, 1385–1388.

Phillippy, B.Q. (2002). Stability of plant and microbial phytases. In Food Phytates. Eds. Reddy, N.R. and Sathe, S.K. CRC Press, Boca Raton, FL, USA. pp. 107–126.

Phillippy, B.Q. and Wyatt, C.J. (2001). Degradation of phytate in foods by phytases in fruits and vegetable extracts. J. Food Sci. *66*, 535–539.

Pilu, R., Panzeri, D., Gavazzi, G., Rasmussen, S.K., Consonni, G. and Nielsen, E. (2003). Phenotypic, genetic and molecular characterization of a maize low phytic acid mutant (lpa241). Theor. Appl. Genet. *107*, 980–987.

Porres, J.M., Aranda, P., López-Jurado, M. and Urbano, G. (2003). Effect of natural and controlled fermentation on chemical composition and nutrient dialyzability from beans (*Phaseolus vulgaris*). J. Agric. Food Chem. *51*, 5144–5149.

Porres, J.M., Aranda, P., López-Jurado, M. and Urbano, G. (2005). Nutritional potential of raw and free α-galactosides lupin (*Lupinus albus* Var. *multolupa*) seed flours. Effect of phytase treatment on nitrogen and mineral dialyzability. J. Agric. Food Chem. *53*, 3088–3094.

Porres, J.M., Etcheverry, P., Miller, D.D. and Lei, X.G. (2001). Phytase and citric acid supplementation in whole-wheat bread improves phytate-phosphorus release and iron dialyzability. J. Food Sci. *66*, 614–619.

Quan, C.S., Fan, S.D. and Ohta, Y. (2003). Immobilization of *Candida krusei* cells producing phytase in alginate cell beads: an application of the preparation of *myo*-inositol phosphates. Appl. Microbiol. Biotechnol. *62*, 41–47.

Quan, C.S., Fan, S.D., Zhang, L.H., Wang, Y.J. and Ohta, Y. (2002). Purification and properties of a phytase from *Candida krusei* WZ-001. J. Biosci. Bioeng. *94*, 419–425.

Quan, C.S., Tian, W.J., Fan, S.D. and Kikuchi, Y. (2004). Purification and properties of a low-molecular-weight phytase from *Cladosporium* sp. FP-1. J. Biosci. Bioeng. *94*, 260–266.

Ravindran, V.B.W.K.E. (1995). Phytates: Occurrence, bioavailibility and implications in poultry nutrition. Poult. Avian. Biol. Rev. *6*, 125–143.

Rodriguez, E., Porres, J.M., Han, Y. and Lei, X.G. (1999a). Different sensitivity of recombinant *Aspergillus niger* phytase (r-PhyA) and *Escherichia coli* pH 2.5 acid phosphatase (r-AppA) to trypsin and pepsin *in vitro*. Arch. Biochem. Biophys. *365*, 262–267.

Rodriguez, E., Han, Y.M. and Lei, X.G. (1999b). Cloning, sequencing and expression of an *Escherichia coli* acid phosphatase/phytase gene (*AppA2*) isolated from pig colon. Biochem. Biophys. Res. Commun. *257*, 117–123.

Rodriguez, E., Mullaney, E.J. and Lei, X.G. (2000a). Expression of the *Aspergillus fumigatus* phytase gene in *Pichia pastoris* and characterization of the recombinant enzyme. Biochem. Biophys. Res. Commun. *268*, 373–378.

Rodriguez, E., Wood, Z.A., Karplus, P.A. and Lei, X.G. (2000b). Site-directed mutagenesis inproves catalytic efficiency and thermostability of *Escherichia coli* pH 2.5 acid phosphatase/phytase expressed in *Pichia pastoris*. Arch. Biochem. Biophys. *382*, 105–112.

Sabu, A., Sarita, S., Pandey, A., Bogar, B., Szakacs, G. and Soccol, C.R. (2002). Solid-state fermentation for production of phytase by *Rhizopus oligosporus*. Appl. Biochem. Biotechnol. *102–103*, 251–260.

Saito, T., Kohno, M., Tsumura, K., Kugimiya, W. and Kito, M. (2001). Novel method using phytase for separating soybean β-conglycinin and glycinin. Biosci. Biotechnol. Biochem. *65*, 884–887.

Sajidan, A., Farouk, A., Greiner, R., Jungblut, P., Müller, E.C. and Borriss, R. (2004). Molecular and physiological characterization of a 3-phytase from soil bacterium *Klebsiella* sp. ASR1. Appl. Microbiol. Biotechnol. *65*, 110–118.

Sandberg, A. and Andlid, T. (2002). Phytogenic and microbial phytases in human nutrition. Int. J. Food Sci. Tech. *37*, 823–833.

Sano, K., Fukuhara, H. and Nakamura, Y. (1999). Phytase of the yeast *Arxula adenivorans*. Biotechnol. Lett. *21*, 33–38.

Schenk, G., Guddat, L.W., Ge, Y., Carrington, L.E., Hume, D.A., Hamilton, J. and de Jersey, J. (2000) Identification of mammalian-like purple acid phosphatases in a wide range of plants. Gene *250*, 117–125.

Schoner, F.J., Hoppe, P.P. and Schwarz, G. (1991). Comparative effects of microbial phytase and inorganic phosphorus on performance and on retentions of phosphorus, calcium and crude ash in broilers. J. Anim. Physiol. Anim. Nutr.-Zeitschrift fur Tierphysiologie Tierernahrung und Futtermittelkunde. *66*, 248–255.

Scott, J.J. (1991). Alkaline phytase activity in nonionic detergent extracts of legume seeds. Plant Physiol. *95*, 1298–1302.

Scott, J.J. and Loewus, F.A. (1986) A calcium-activated phytase from pollen of *Lilium longiflorum*. Plant Physiol. *82*, 333–335.

Segueilha, L., Lambrechts, C., Boze, H., Moulin, G. and Galzy, P. (1992). Purification and properties of the phytase from *Schwanniomyces castellii*. J. Ferment. Bioeng. *74*, 7–11.

Shin, S., Ha, N.C., Oh, B.C., Oh, T.K. and Oh, B.H. (2001). Enzyme mechanism and catalytic property of β propeller phytase. Structure *9*, 851–858.

Simon, O. and Igbasan, F. (2002). *In vitro* properties of phytases from various microbial origins. Int. J. Food Sci. Technol. *37*, 813–822.

Skoglund, E., Larsen, T. and Sandberg, A.S. (1997). Comparison between steeping and pelleting a mixed diet at different calcium levels on phytate degradation in pigs. Can. J. Anim. Sci. *77*, 471–477.

Stahl, C.H., Wilson, D.B. and Lei, X.G. (2003). Comparison of extracellular *Escherichia coli* AppA phytases expressed in *Streptomyces lividans* and *Pichia pastoris*. Biotechnol. Lett. *25*, 827–831.

Stevenson-Paulik, J. Bastidas, R.J., Chiou, S.T., Frye, R.A. and York, J.D. (2005). Generation of phytate-free seeds in Arabidopsis through disruption of inositol polyphosphate kinases. Proc. Natl. Acad. Sci. USA. *102*, 12612–12617.

Stöckmann, C., Losen, M., Dahlems, U., Knocke, C., Gellisen, G. and Buchs, J. (2003). Effect of oxygen supply on passaging, stabilising, and screening of recombinant *Hansenula polymorpha* production strains in test tube cultures. FEMS. Yeast. Res. *4*, 195–205.

Tomschy, A., Brugger, R., Lehmann, M., Svendsen, A., Vogel, K., Kostrewa, D., Lassen, S.F., Burger, D., Kronenberger, A., van Loon, A.P.G.M., Pasamontes, L. and Wyss, M. (2002). Engineering of phytase for improved activity at low pH. Appl. Environ. Microbiol. *68*, 1907–1913.

Tomschy, A., Tessier, M., Wyss, M., Brugger, R., Broger, C., Schnoebelen, L., van Loon, A.P.G.M. and Pasamontes, L. (2000). Optimization of the catalytic properties of *Aspergillus fumigatus* phytase based on the three-dimensional structure. Prot. Sci. *9*, 1304–1311.

Türk, M., Sandberg, A.S., Carlsson, N. and Andlid, T. (2000). Inositol hexaphosphate hydrolysis from baker's yeast. Capacity, kinetics, and degradation products. J. Agric. Food Chem. *48*, 100–104.

Ullah, A.H.J. and Cummins, B.J. (1987) Purification, N-terminal amino acid sequence and characterization of pH 2.5 optimum acid phosphatase (E.C.3.1.3.2) from *Aspergillus ficuum*. Prep. Biochem. *17*, 397–422.

Ullah, A.H.J. and Cummins, B.J. (1988) *Aspergillus ficuum* extracellular pH 6.0 optimum acid phosphatase: purification, N-terminal amino acid sequence, and biochemical characterization. Prep. Biochem. *18*, 37–65.

Ullah, A.H.J. and Dischinger Jr., H.C. (1993) *Aspergillus ficuum* phytase: complete primary structure elucidation by chemical sequencing. Biochem. Biophys. Res. Commun. *92*, 747–753.

Van der Kaay, J. and Van Haastert, P.J.M. (1995). Stereospecificity of inositol hexakisphosphate dephosphorylation by *Paramecium* phytase. Biochem. J. *312*, 907–910.

Van Etten, R.L., Davidson, R., Stevis, P.E., MacArthur, H. and Moore, D.L. (1991) Covalent structure, disulfide bonding, and identification of reactive surface and active site residues of human prostatic acid phosphatase. J. Biol. Chem. *266*, 2313–2319.

Vieira, E.C. and Nogueira, A.R.A. (2004). Orthophosphate, phytate, and total phosphorus determination in cereals by flow injection analysis. J. Agric. Food Chem. *52*, 1800–1803.

Viveros, A., Centeno, C., Brenes, A., Canales, R. and Lozano, A. (2000). Phytase and acid phosphatase activities in plant feedstuffs. J. Agric. Food Chem. *48*, 4009–4013.

Vohra, A. and Satyanarayana, T. (2001). Phytase production by the yeast, *Pichia anomala*. Biotechnol. Lett. *23*, 551–554.

Wang, X.Y., Meang, F.G. and Zhou, H.M. (2004) The role of disulfide bonds in the conformational stability and catalytic activity of phytase. Biochem. Cell Biol. *82*, 329–334.

Wise, A. and Gilburt, D.J. (1982). Phytate hydrolysis in germfree and conventional rats. Appl. Environ. Microbiol. *43*, 753–756.

Wodzinski, R.J. and Ullah, A.H.J. (1996) Phytase. Adv. Appl. Microbiol. *42*, 263–302.

Wyss, M., Brugger, R., Kronenberger, A., Rémy, R., Fimbel, R., Oesterhelt, G., Lehmann, M. and van Loon, A.P.G.M. (1999a). Biochemical characterization of fungal phytases (*myo*-inositol hexakisphosphate phosphohydrolases): Catalytic properties. Appl. Environ. Microbiol. *65*, 367–373.

Wyss, M., Pasamontes, L., Friedlein, A., Rémy, R., Tessier, M., Kronenberger, A., Middendorf, A., Lehmann, A., Scnoebelen, L., Röthlisberger, U., Kusznir, E., Wahl, G., Müller, F., Lahm, H.W., Vogel, K. and van Loon, A.P.G.M. (1999b) Biophysical characterization of fungal phytases (*myo*-inositol hexakisphosphate phosphohydrolases): molecular size, glycosylation pattern, and engineering of proteolytic resistance. Appl. Environ. Microbiol. *65*, 359–366.

Wyss, M., Pasamontes, L., Rémy, R., Kohler, J., Kusznir, E., Gadient, M., Müller, F. and van Loon, A.P.G.M. (1998). Comparison of the thermostability properties of three acid phosphatases from molds: *Aspergillus fumigatus* phytase, *A. niger* phytase, and *A. niger* pH 2.5 acid phosphatase. Appl. Environ. Microbiol. *64*, 4446–4451.

Yanke, L.J., Bae, H.D., Selinger, L.B. and Cheng, K.J. (1998). Phytase activity of anaerobic ruminal bacteria. Microbiol. *144*, 1565–1573.

Yanke, L.J., Selinger, L.B. and Cheng, K.J. (1999). Phytase activity of *Selenomonas ruminantium*: a preliminary characterization. Lett. Appl. Microbiol. *29*, 20–25.

Yi, Z., Kornegay, E.T., Ravindran, V. and Denbow, D.M. (1996). Improving phytate phosphorus availability in corn and soybean meal for broilers using microbial phytase and calculation of phosphorus equivalency values for phytase. Poult. Sci. *75*, 240–249.

Yin, C., Zhu, Z.Z., Lin, X.A., Yi, Y.Z., Zhang, Z.F. and Shen G.F. (2005). Overexpression and characterization of *appA* phytase expressed by recombinant baculovirus-infected silkworm. J. Microbiol. Biotechnol. *15*, 466–471.

Zhang, Z.B., Kornegay, E.T., Radcliffe, J.S., Denbow, D.M., Veit, H.P. and Larsen, C.T. (2000a). Comparison of genetically engineered microbial and plant phytase for young broilers. Poult. Sci. *79*, 709–717.

Zhang, Z.B., Kornegay, E.T., Radcliffe, J.S., Wilson, J.H. and Veit, H.P. (2000b). Comparison of phytase from genetically engineered *Aspergillus* and canola in weanling pig diets. J. Anim. Sci. *78*, 2868–2878.

Zinin, N.V., Serkina, A.V., Gelfand, M.S., Shevelev, A.B. and Sinepky, S.P. (2004). Gene cloning, expression and characterization of novel phytase from *Obesumbacterium proteus*. FEMS. Microbiol. Lett. *236*, 283–290.

CHAPTER 30

NITRILE HYDROLASES

PRAVEEN KAUL, ANIRBAN BANERJEE AND UTTAM CHAND BANERJEE*

Department of Pharmaceutical Technology (Biotechnology), National Institute of Pharmaceutical Education and Research, SAS Nagar, Punjab, India
**ucbanerjee@niper.ac.in*

1. INTRODUCTION

Nitrile compounds (chemical formula RCN) are widespread in the environment. In nature they are mainly present as cyanoglycosides which are produced by plants and animals. Plants also produce other nitrile compounds such as cyanolipids, ricinine and phenylacetonitrile. Chemical industries use various nitrile compounds extensively for manufacturing a variety of polymers and other chemicals. From the synthetic viewpoint, nitriles represent a widely applicable chiral synthon which can be employed for the homologation of a carbon framework. Further transformations of the nitriles thus obtained are impeded due to the harsh reaction conditions required for its hydrolysis. In this context, the *chemo*-selective biocatalytic hydrolysis of nitriles represents a valuable alternative because it occurs at ambient temperature and near physiological pH (Banerjee *et al*, 2002). Most the nitriles are highly toxic, mutagenic and carcinogenic. The general toxicities of nitriles in humans are expressed as gastric diseases and vomiting (nausea), bronchial irritation, respiratory distress, convulsions, coma and osteolathrysm which leads to lameness and skeletal deformities. Nitriles inactivate the respiratory system by tightly binding to cytochrome-c-oxidase (Solomonson and Spehar, 1981). Despite their toxicity, large quantities of cyanides are used in the metal plating, pharmaceutical, agricultural and chemical industries, thus they are widely distributed in industrial wastewater. Currently used chemical methods for detoxification of cyanides such as alkaline chlorination, ozonization and wet-air oxidation are expensive and require the use of hazardous chemicals such as chlorine and sodium hypochlorite. Microbial degradation has been considered an efficient way of removing highly toxic nitriles from the environment. Biological methods are more acceptable

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 531-547

and environmentally friendly than chemical methods. In recent years the use of nitrile-converting enzymes, especially nitrilase and nitrile hydratase, have increased tremendously.

2. THE NITRILASE SUPERFAMILY

On the basis of structure and sequence analyses a new family of enzymes, termed the nitrilase superfamily, was constructed (Brenner, 2002). The nitrilase superfamily consists of 13 branches but members of only one branch are known to have true nitrilase activity, whereas 8 or more branches have apparent amidase or amide condensation activities. Nitrilases are found in plants, animals and fungi, and many of these organisms have multiple nitrilase-related proteins from different branches of the superfamily. Related sequences are also found in phylogenetically distinct prokaryotes. These observations suggest that the superfamily emerged prior to the separation of plants, animals and fungi and then spread laterally to bacteria and archea. Due to the historical observation that aliphatic amidases are related to nitrilases, the superfamily was designated as the nitrilase superfamily. Nitrile hydratases (NHases), amidase signature enzymes and thiol proteases are not included in this superfamily. All the superfamily members contain a conserved apparently catalytic triad of glutamate, lysine and cysteine, and a largely similar α-β-β-α architecture.

2.1. Nitrilase

Nitrilase, the first nitrile-hydrolysing enzyme described some 40 years ago, was known to convert indole 3-acetonitrile to indole 3-acetic acid (Thimann and Mahadevan, 1964). More recently, the biotechnological potential of nitrile hydrolysing enzymes and especially nitrilases (O'Reilly and Turner, 2003) has led to the isolation of a range of bacteria and fungi capable of hydrolysing nitriles. The enzymes show a diverse range of biochemical characteristics. In particular, the substrate specificities of the enzymes vary widely. Initial investigations suggested them to be specific for aromatic nitriles. However, this distinction was reconsidered on the light of growing information on this class of enzymes. Nitrilases are now reported from different sources which are active on aliphatic as well as on heterocyclic nitriles. Broadly these enzymes are classified into three different categories based on their substrate specificities, namely aliphatic, aromatic and arylaceto-nitrilases. These are generally inducible enzymes and do not require metal co-factors or prosthetic groups for their activity. For the development of a biocatalytic process a library of enzymes (micro-organisms) has to be screened. Screening is generally carried out using a robust assay method which is highly specific for the targeted enzymatic system, has a broad linearity range, is able to detect the product of interest and is rapid to perform and therefore amenable to high-throughput. The nitrilase assay generally used is based on measurement of the equimolar amount of ammonia (by-product) produced. Whilst nitrilase

activity can be determined by direct estimation of the carboxylic acid formed, either by HPLC (Lebuhn and Hartmann, 1993) or by infrared spectroscopy (FTIR) (Dadd *et al*, 2000), these methods have certain serious drawbacks. HPLC is not high-throughput amenable because of the long time requirement for sample preparation, and the FTIR methodology - whilst providing real-time kinetic analysis of nitrile biocatalysis – is limited by the availability of the type of instrument and silicon probe necessary thereby reducing the practical usefulness of the technique. Thus ammonia estimation, though indirect, is the more practical route for the determination of nitrilase activity. Currently available methods for ammonia estimation include ninhydrin (Khachadurian *et al*, 1960), Nesslerization and the Berthelot method (Fawcett and Scott, 1960). The ninhydrin and Nesslerization methods suffer from several disadvantages including the requirement for large sample volumes because of low sensitivity and interference by organic solvents and inorganic ions (Kruse and Mellon, 1953). Although the Berthelot method appears to be quite sensitive (20 to 200 μM), it involves the use of corrosive reagents like phenol for the generation of the chromophore. The method also requires heating of the test solution at 90 °C for 30 min or at 100 °C for 5 min because at room temperature at least 2 h is required for the development of a stable chromophore. Heating is disadvantageous because it increases the production of toxic phenol vapours and generates insoluble MnO_2 (where $MnSO_4$ is used as a catalyst) which interfers with the spectrophotometric determination (Krallmann-Wenzel, 1985). To overcome the disadvantages associated with these existing techniques, an attempt was made to develop a rapid, simple and sensitive method for the estimation of nitrilase activity (Banerjee *et al*, 2003a). Ammonia liberated due to nitrilase activity was reacted with *o*-phthaldialdehyde and 2-mercaptoethanol to form an isoindole derivative which fluoresced at $\lambda_{max(emiss)}$ 467 nm when excited at $\lambda_{max(excit)}$ 412 nm. Primary amines also react with *o*-phthaldialdehyde in the presence of thiols to produce fluorescent 1-alkyl (and aryl) thio-2-alkylisoindoles (Simons and Johnson, 1976). In the absence of thiol compound *o*-phthaldialdehyde reacts with ammonia to produce a non-fluorescent product (Simons and Johnson, 1978). The flourimetric method gives more accurate results at higher substrate concentrations and may therefore be useful for screening purposes. The design of a robust screening method involves the use of high concentrations of substrate in the reaction mixture which is desirable from the application point of view. This method may therefore be more useful during the early stages of screening to search for more potent enzymes or micro-organisms.

Analytical methods generally employed for screening for an enantioselective biocatalyst include HPLC, GC, LC-MS *etc*., which are not readily adaptable to high-throughput. Assaying enzyme catalysed transformations in a high-throughput format is crucial to enzyme discovery and enzyme engineering. The lack of a suitable plate assay method for nitrile hydrolyzing enzymes renders the screening step as rate limiting in the search for micro-organisms having the desired activity. The intracellular nature of this class of enzymes creates problems regarding the development of a solid agar plate assay. However, assaying liquid medium (reaction

mixture) for acid productioncan be developed into a high throughput format for the screen. A simple, colorimetric, pH responsive method was developed for the rapid enantioselective screening of nitrile hydrolysing enzyme libraries (Banerjee *et al*, 2003b). pH sensitive indicators have been used to monitor various enzyme catalysed reactions such as carbonic anhydrase (Gibbons and Edsall, 1963), cholinesterase (Lowry *et al*, 1954), lipase (Baumann *et al*, 2000) *etc*. The method is based on the drop in pH that occurs as the reaction proceeds due to the formation of acid. The pH drop is reflected by the colour change of the indicator, provided the colour profile of the indicator falls within the pH range of the enzyme activity. For the colour change to be proportional to the number of protons released, both the indicator and the buffer must have similar affinities for protons (pK_as within 0.1 unit of each other), such that the relative amounts of protonated buffer to protonated indicator remain constant during the course of the reaction. This method was utilized to screen nitrilase producing micro-organisms for the production of (*R*)-(−)-mandelic acid, a versatile chiral building block (Kaul *et al*, 2004b). Since only the (*R*)-isomer of mandelonitrile is commercially available, enantioselectivity was determined by comparing the rate of colour turnover of the (*R*)-isomer with that of the racemate. Conclusive statements regarding the enantioselectivity of the micro-organisms could not be made in cases where both changed colour simultaneously, reflecting the importance of using pure enantiomers. The proposed method is simple and, most importantly, less time consuming. Using this pH responsive strategy a large number of micro-organisms can be analysed simultaneously in a short time, thus reducing the number of samples to be analysed in greater quantitative detail (HPLC analysis). The assay requires very little substrate thus allowing the use of pure enantiomers which are not always available in large quantities. The disadvantage of the method is its qualitative nature. The presence of cells interferes with the spectrophotometric reading. The use of cell free extracts may solve the problem but will also add an extra step to the screening procedure. Moreover, this may require special instrumentation (*e.g.* a microplate reader) which can be avoided by visualising the enzyme catalysed reaction using a suitable indicator. Hence, by compromising quantitative aspects screen throughput was increased. The goal was to use the method not for precise quantitation but for screening large numbers of micro-organisms to facilitate the identification of those having the desired enantioselectivity acceptable for the development of a biocatalytic resolution process.

2.1.1. Distribution and physiological role of nitrilases

Nitrilases are found relatively infrequently in nature. The existence of the enzyme activity in 3 out of 21 plant families (*Gramineae, Cruciferae* and *Musaceae*) (Thimann and Mahadevan, 1964) and in a limited number of fungal genera (*Fusarium, Aspergillus, Penicillium*) (Harper, 1977) indicates the relative scarcity of this activity. Nitrile-degrading activity is more frequent in bacteria, though without extensive screening it is almost impossible to assess the actual distribution frequency. A number of bacteria (*Acinetobacter, Corynebacterium, Arthrobacter, Pseudomonas, Klebsiella, Nocardia, Rhodococcus etc*.) are known to utilize nitriles

as sole sources of carbon and nitrogen. The physiological role of nitrile hydrolysing enzymes in micro-organisms is not clear. In plants such activities are implicated in nutrient metabolism, particularly in the degradation of glucosinolates (Bestwick *et al*., 1993) and in the synthesis of indole acetic acid (Bartel and Fink, 1994).The genome sequence of *Arabidopsis thaliana* revealed four nitrilase-related sequences (AtNIT 1, 2, 3 and 4) (Bartling *et al*., 1992). AtNIT1, 2 and 3 are isoenzymes which are found only in *Brassicaceae* (Hillebrand *et al*., 1998). AtNIT4 on the other hand is not related to the other three nitrilases and its homologues are found in many plant species such as tobacco (Tsunoda and Yamaguchi, 1995) and rice (Piotrowski *et al*., 2001). The major physiological role for these enzymes appears to be glucosinolate metabolism. AtNIT4 is a β-cyano alanine hydratase and plays a role in cyanide detoxification (Piotrowski *et al*., 2001). In microbial systems nitrilases probably form components of a complex pathway controlling the production and degradation of aldoximes. Nitriles, which are formed by enzyme activities upstream of aldoxime dehydratases, may further undergo hydrolysis, oxidation, reduction *etc*. by different enzymes including nitrilases (Fig. 1).

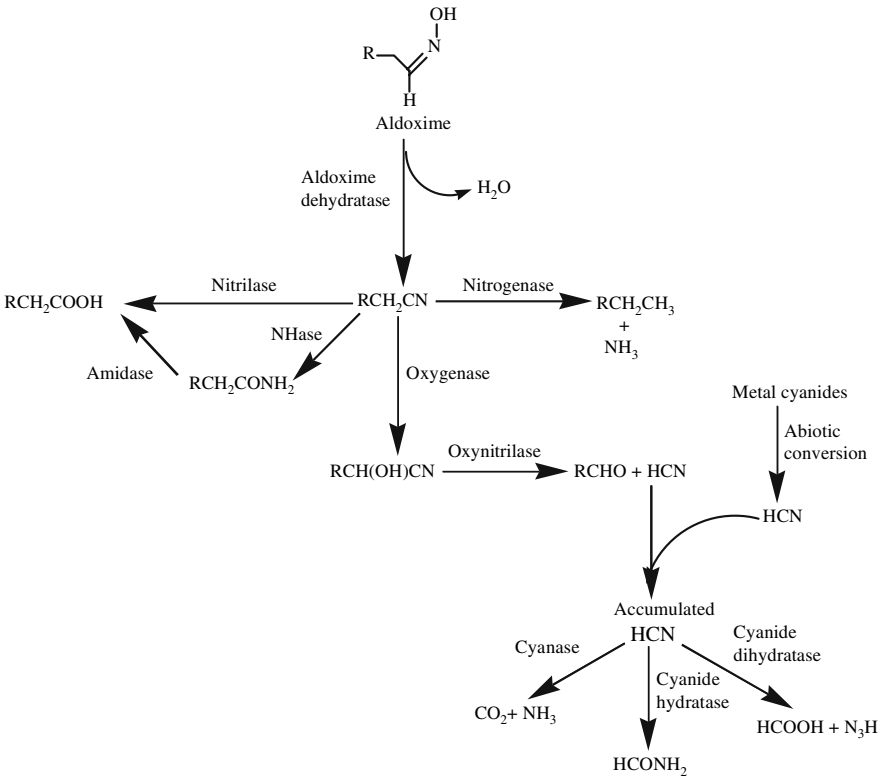


Figure 1. Different pathways of nitrile metabolism

2.1.2. Molecular structure and genetic analysis

To date many nitrilases have been purified and their subunit molecular mass and primary structures have been determined. Most nitrilases consist of a single polypeptide having a molecular mass in the range 30 to 45 kDa, which aggregate to form the active holoenzyme under different conditions. The preferred form of the enzyme seems to be a large aggregate of 6 to 26 subunits. Most of the enzymes show substrate dependent activation, though the presence of elevated concentrations of salt, organic solvents, pH, temperature or even the enzyme itself may also trigger subunit association and therefore activation (Nagasawa *et al*., 2000). The hydrophobic effect resulting from the above mentioned conditions might change the conformation of the enzyme thereby exhibiting hydrophobic sites and enabling subunit assembly and enzyme activation. The nitrilases of *Pseudomonas fluorescens* DSM 7155 and *Bacillus pallidus* Dac 521 were co-purified with chaperonins (Almatawah *et al*., 1999; Layh *et al*., 1998). It was suggested that chaperonins might play a role in the assembly of the subunits into high molecular weight complexes and also their stabilisation. When the nitrilase gene of *Comamonas testosteroni* was cloned and over-expressed in *E. coli*, as much as 30% of the total cellular protein was found to be nitrilase but only 10% of this was soluble in nature. The solubility of nitrilase was enhanced by co-overexpression of GroEL (chaperonin) in the same cell and the activity increased five-fold (Schill *et al*., 1995). This further confirms the crucial role of chaperonins in the correct folding and subunit association of nitrilase.

The genes encoding the nitrilases of *Klebsiella pneumoniae* ssp. *ozaenae* (Stalker *et al*., 1988), *R. rhodochrous* strain K22 (Kobayashi *et al*., 1992), *A. faecalis* JM3 (Kobayashi *et al*., 1993), *C. testosteroni* (Schill *et al*., 1995) and *Bacillus* sp. strain OxB-1 (Kato *et al*., 2000) have been sequenced and over-expressed. The four nitrilases of *A. thaliana* have also been over-expressed in *E. coli*. AtNIT1, 2 and 3 are clustered in a 13.8 kb region on chromosome 3 and share 80% identity in sequences (Hillebrand *et al*., 1998). AtNIT4, not linked to these genes, shows 63% identity with AtNIT1-3, which reflects their different functionality in *A. thaliana* (Piotrowski *et al*., 2001). A Cys residue has been shown to be essential for the catalytic activity of the nitrilases (Kobayashi *et al*., 1992, 1993). This Cys is proposed to be a part of the catalytic triad Cys-Glu-Lys which has been identified from the crystal structures of the NitFhit and N-carbamyl-D-amino acid amidohydrolases, members of nitrilase superfamily (Nakai *et al*., 2000). The genetic regulation of nitrilases is largely unknown. A gene (nitR) was identified downstream of the *R. rhodochrous* J1 nitrilase gene and encodes a protein that has significant homology to AraC type transcriptional regulators (Martin and Rosner, 2001). NitR appears to be a positive regulator of nitrilase gene (nitA) expression (Komeda *et al*., 1996a). Unlike most AraC type regulators whose genes are transcribed divergently from the genes they are regulating, transcriptional analysis of nitR indicates that it is probably expressed by read through from the nitA promoter (Komeda *et al*., 1996b). A putative nitrilase regulator has also been found in *Bacillus* sp. strain OxB-1 (Kato *et al*., 2000).

2.1.3. *Reaction mechanism*

Hydrolysis of a nitrile of the form R-C ≡ N produces the corresponding amide, R-C = O(NH_2), with the addition of one molecule of water and the corresponding acid, R-CO_2H, with the addition of second molecule of water. Nitrilases are interesting enzymes in the sense that the substrates are nitriles but the reaction does not involve the release of or the reaction with a substantial amount of the corresponding amide. However, it has been shown that the purified nitrilases of *F. oxysporum* ssp. *melonis* (Goldhust and Bohak, 1989), *Pseudomonas fluorescens* DSM7155 (Layh *et al*, 1998) and the ricinine nitrilase of *Pseuodomonas* sp. (Hook and Robinson, 1964) produce a small amount of amide product and therefore have nitrile hydratase activity. In these cases the amide product is usually < 5% of the total reaction products. AtNIT1 and AtNIT4 also have NHase activity, especially AtNIT4 which produces 1.5 times more asparagine than aspartic acid from β-cyano alanine (Piotrowski *et al*, 2001). Some of the nitrilases use amides, albeit at a very low rate compared to the nitrile substrate, and therefore have amidase activity (Kobayashi *et al*, 1998). A reaction mechanism was proposed which accounts for all these activities (Hook and Robinson, 1964; Thimann and Mahadevan, 1964). This mechanism involves a nucleophilic attack on the nitrile carbon by the sulfhydryl group of the nitrilase, leading to a tetrahedral intermediate via an enzyme thioimidate route (Fig. 2). The tetrahedral intermediate can frequently break down anomalously to produce an amide instead of the normal acid product. Nitrilase interaction with its substrates and intermediates has some geometric constraints: linear substrate (approximately 180 °C), planar

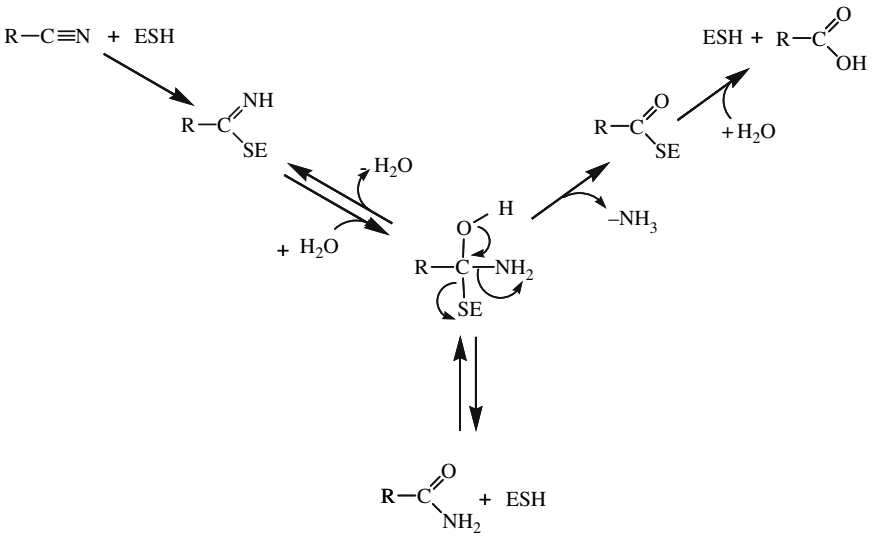


Figure 2. Reaction mechanism of nitrilase catalysis

thioimidate and acylenzyme intermediates (approximately 120 °C), and tetrahedral water-bonded intermediates (approximately 109.5 °C). Since ammonia is a better leaving group than the enzyme, the first enzyme-dependent water addition does not produce an acid amide but rather an acylenzyme complex the hydrolysis of which produces the acid product. Most nitrilases bind strongly to a bulky substrate R group in a conformation that places the second carbon closer to 120° than to 180° from the cyano nitrogen. Hence, fitting a distorted nitrile substrate would drive the reaction towards thioimidation rather than tetrahedral intermediate. In support of this view, most nitrilases are noted to prefer bulky substrates to non-substituted acetonitrile (Pace and Brenner, 2001).

2.2. Nitrile Hydratase

Nitrile hydratase (NHase) is a key enzyme in the bienzymatic pathway of the conversion of nitriles to amides, which are further converted to the corresponding acid by amidases. A number of micro-organisms having NHase activity have been isolated and the enzymes have been purified and characterized. These revealed the wide ranging physiochemical properties and substrate specificities of the NHases, which are composed of two types of subunits (α and β) complexed in varying numbers. They are metalloenzymes containing either cobalt or iron. On the basis of the metal ion present, the NHases can be classified into two broad groups, namely ferric NHases and cobalt NHases.

2.2.1. Ferric NHases

The NHases from *Rhodococcus* R312 (formerly known as *Brevibacterium* R312) and *Pseudomonas chlororaphis* B23 are the first examples of non-heme iron enzymes containing a low spin Fe (III) ion (Sugiura *et al*., 1987). These have been well characterized by ESR, extended X-ray absorption (EXFAS), electron nuclear double resonance (ENDOR) and Raman resonance spectroscopies. These studies revealed that the NHase from *Rhodococcus* sp. R312 is a $(\alpha\beta)_2$ tetramer containing two low spin non-heme ferric ions which exist in a tetragonally distorted octahedral ligand field of three histidine imidazoles, two cysteine thiolates and a hydroxide. The activity of NHase has unique features when exposed to light (Endo *et al*., 1999). During aerobic incubation in the dark NHase activity in *Rhodococcus* sp. N-771 decreases, but this activity is almost completely recovered upon irradiation with visible light. This photo-reactivity of NHase is intrinsic to the enzyme, and the mechanism has been biochemically elucidated (Endo *et al*., 1999). The chromophore involved in the photo-activation is an iron complex in the β subunit, and light irradiation of the complex induces a conformational change of the subunit. Moreover, activity is regulated by nitric oxide (NO). An endogenous NO molecule is bound to the non-heme iron (III) centre in the inactive NHase and is released upon photo-activation, resulting in recovery of the original NHase activity. Three-dimensional analysis of *Rhodococcus* R312 NHase showed that it contains a novel iron centre (Huang *et al*., 1997). All the metal ion protein ligands are contained

within the α subunit. Three cysteine thilotes and two main chain nitrogen atoms are ligands. These five iron ligands (Cys 110, Cys113, Cys115, Ser 114 and Cys 115) are located on five vertices of an octahedron, the sixth position being unoccupied. Considering the experimental data for hydrogen and oxygen ENDOR resonances, a hydroxide ion is likely to occupy this position. Furthermore, Cys112 and Cys114 (corresponding to Cys113 and Cys115 in the *Rhodococus* R312 NHase) are post-translationaly oxidised to cysteine-sulfinic and sulfenic acids, respectively, in the *Rhodococcus* N-771 NHase.

2.2.2. Cobalt NHase

In the presence of cobalt ions the actinomycete *R. rhodochrous* J1 (Komeda *et al*, 1996b) produces two Nhases, depending on the inducer. When cultured in a medium containing urea and cyclohexane carboxamide, high and low molecular weight NHases (H- and L-NHases) are selectively induced (Yamada and Kobayashi, 1996). H-NHases (Nagasawa *et al*, 1991) act preferentially on aliphatic nitriles, whereas L-NHases (Komeda *et al*, 1996b) exhibit higher affinity for aromatic nitriles. H- and L-NHases have been used for the industrial production of acrylamide and nicotinamide from acrylonitrile and 3-cyanopyridine, respectively. Both purified NHases contain cobalt as co-factor. The cobalt in H-NHase exists as a low-spin Co ion in a tetragonally distorted octahedral ligand field. The similarities of the pre-edge and extended X-ray absorption fine structure (EXAFS) spectra suggests that the ligand environments of the metal ions in the cobalt and iron containing NHases are similar. Recently, a NHase from *Pseudomonas putida* with stereoselectivity toward 2-(S)-(4-chlorophenyl)-3-methylobutyronitrile was found to contain noncorrin low-spin Co (Payne *et al*, 1997). The three dimensional structures of both the *Rhodococcus rhodochrous* J1 and *Pseudomonas putida* NHase are probably similar to that of the *Rhodococcus* R 312 NHase because of their sequence similarities. The cobalt NHases have threonine in the –V-C-(T/S)-L-C-S-C-sequence as the active site, (Payne *et al*, 1997) whereas the ferric NHase has serine. The difference in the metal co-factors may be ascribed to the different amino acid residues at this position. Spectroscopic and structural analyses of novel NHases in *Agrobacterium tumifaciens* which contain cobalt and iron should improve our understanding of the functions of these metals. There may be two main reasons for having Co in the active site: 1) Cobalt is a very good catalyst for CN-triple-bond hydration, and 2) it is required for the folding of the enzyme. In addition to the function of the active centre, cobalt ions may play a role in enhancing the folding or stabilization of the subunit polypeptides of the enzyme.

2.2.3. Reaction mechanism

One possible reaction model is that catalysis proceeds without direct coordination of the substrate to the metal ion, where the metal ion plays a role as a lewis acid activating a water molecule. The next step could be either of the following: (1) the nitrile substrate approaches a metal-bound hydroxide ion which can act as a nucleophile attacking the nitrile carbon atom (mechanism I), or (2) a metal bound

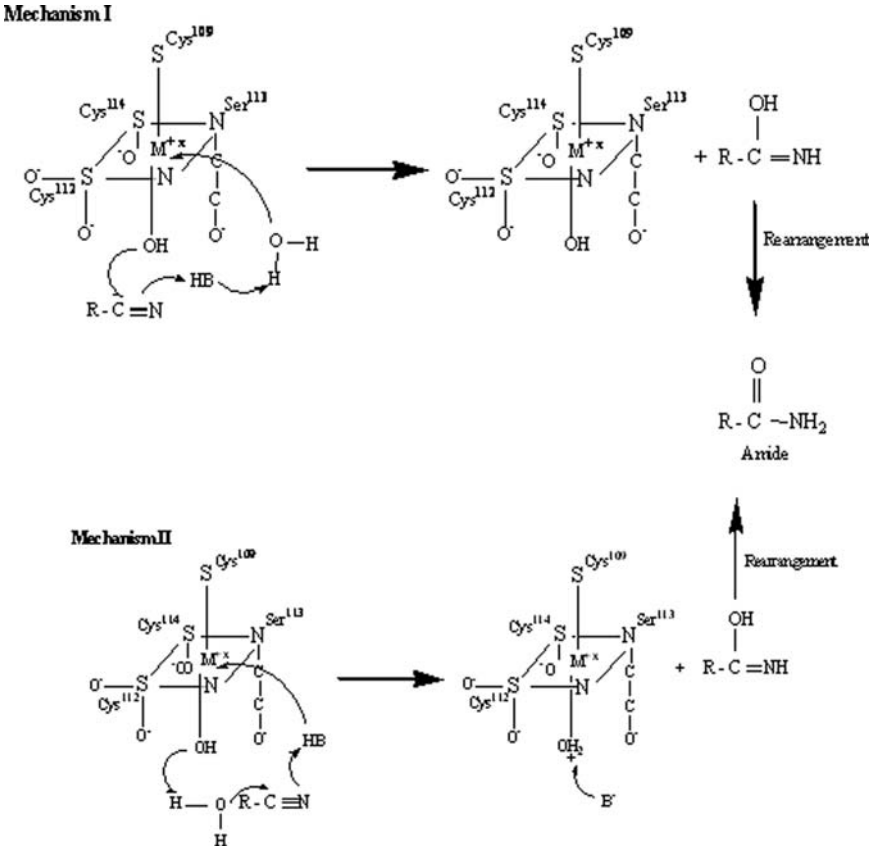


Figure 3. Reaction mechanism of nitrile hydratase catalysed hydrolysis of nitriles

hydroxide ion as a general base activates a water molecule which then attacks the nitrile carbon (mechanism II) resulting in the formation of an imidate. The imidate finally tautomerizes to form an amide (Fig. 3).

2.3. Amidase

Amidases catalyse the hydrolysis of carboxylic acid amides to free carboxylic acids and ammonia. These enzymes are involved in nitrogen metabolism in cells and are widely distributed in nature. They have been found both in prokaryotic and eukaryotic cells. Unlike nitrile hydratases, the association of amidases with metals such as cobalt or iron is reported only in case of *K. pneumoniae* (Nawaz *et al*., 1996). Using site-directed mutagenesis it was confirmed that these are sulfhydryl enzymes. The other amidases from *Pseudomonas chlororaphis* B23, *R. erythropolis* MP50, *R. rhodochrous* J1 *etc*. belong to a group of amidases containing the Gly-Gly-Ser-Ser

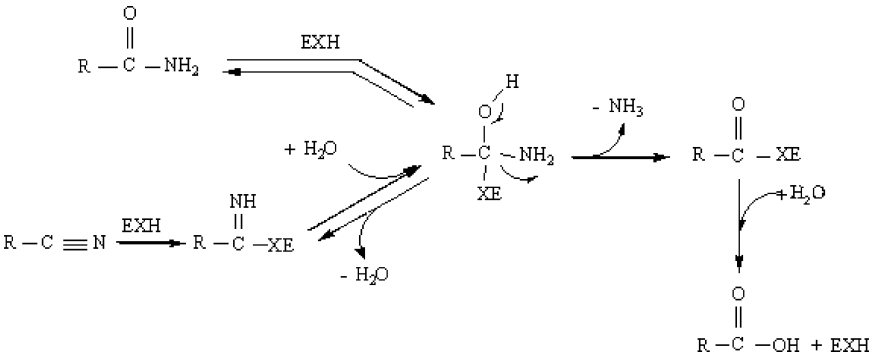


Figure 4. Reaction mechanism of amidase

signature (Chebrou *et al*, 1996). Surprisingly the amidase of *R. rhodochrous* J1 is found to catalyse the hydrolytic cleavage of the nitrile functional group to an acid and ammonia stoichiometrically. The amidase exhibited a K_m of 3.26 mM for benzonitrile in contrast to that of 0.15 mM for benzamide as the original substrate (Kobayashi *et al*, 1998).Thus the reaction mechanisms of both the nitrilase- and amidase-catalysed reactions are analogous, but the active nucleophile present in both enzymes differs. Fig. 4 shows the reaction mechanism of amidase, which also involves nitrile hydrolysis. The carbonyl group of the amide undergoes a nucleophilic attack resulting in the formation of a tetrahedral intermediate which is converted to an acyl-enzyme with the removal of ammonia and subsequently hydrolysed to acid. Although there is no homology between the amidase and nitrilase families, comparison of their reaction mechanisms can provide novel insights for the construction of catalysts required for the hydrolysis of nitriles as well as amides. All of the different amidases also exhibit acyl transfer activity in the presence of hydroxylamine (Fournand *et al*, 1998). Hydroxamic acids thus produced are known to possess good chelating properties and can act as potent inhibitors of matrix metalloproteases (Cawston, 1996). They also have the potential to be used as anti-HIV agents or antimalarial agents (Gao *et al*, 1995).

3. APPLICATIONS

Biocatalysis offers a new dimension, innovative approaches and enormous opportunity to prepare useful chiral compounds (Patel, 2001). Research over the past decade has shown that there are very few barriers to the use of enzymes and/or whole cells as biocatalysts in organic synthesis (Faber, 1997). Enzymes are remarkable catalysts capable of accepting a wide array of complex molecules as substrates and are exquisitely selective, catalysing reactions with unparallel *enantio*- and *regio*-selectivities (Schmid *et al*, 2001; Kaul *et al*, 2004a). Enzymatic transformation of nitriles in particular provides great potential for synthetic chemists. The ability of the enzyme system to convert cyano functionality to either an acid

or an amide is, in itself, of great use. Traditional chemical methods for conversion of nitriles to acids or amides have several drawbacks: (1) reactions need to be carried out either in strongly acidic (6 M HCl) or basic (2 M NaOH) reflux conditions, (2) a requirement for high reaction temperatures, (3) the formation of by-products such as toxic HCN or large amounts of salts, *etc*. Biocatalytic hydrolysis of nitriles is attractive due to its ability to effect reactions in a more 'green' manner and because of the potential for carrying out *chemo-, regio-*, and *enantio*-selectiv transformations.

3.1. Synthetic Biocatalysis

Industrial scale processes for the production of acrylamide (Mitsubishi Rayon Co. Japan) (Nagasawa, 1993; Yamada and Nagasawa, 1994) and nicotinamide (Lonza AG, Switzerland) (Chassin, 1996) involving nitrile converting enzymes are the two most prominent examples of biocatalytic reactions being implemented at larger scale and pave the way for the development of other processes involving this class of enzymes. Nitrilases can also selectively convert a single cyano group of a polynitrile which is virtually impossible using conventional chemical methods. *R. rhodochrous* K22 catalyses the conversion of adiponitrile to 5-cyanovaleric acid, which is an intermediate for the synthesis of nylon-6 (Godtfredsen *et al*., 1985) and Tranexamic acid, a homeostatic drug, is obtained by selective mono-hydrolysis of *trans* 1,4-dicyano cyclohexane by *Acremonium* sp. (Nishise *et al*., 1987). Nitrilases also offer the possibility of stereoselective transformation which includes the production of several important fine chemicals and pharmaceutical intermediates, for example (*S*)-phenyllactic acid (Hashimoto *et al*., 1996), L-α-amino acids (Bhalla *et al*., 1992), $(R)-\beta$-hydroxy acids (Wu and Li, 2003), (*S*)-ibuprofen (Yamamoto *et al*., 1990) *etc*. Although in recent years several applications of nitrilases have been recognized (Kobayashi and Shimizu, 1994; Sugai *et al*., 1997) to date there exist only two industrial scale processes for the production of 1,5-dimethyl 2-piperidone (DuPont, USA) and (*R*)-(−)-mandelic acid (BASF, Germany; Mitsubishi Rayon, Japan) employing these enzymes as biocatalysts. 1,5-dimethyl 2-piperidone (Xolvone™) which has desirable properties for electronics, coatings and solvent applications, is produced from 4-cyanopentanoic acid, which in turn is generated from *regio*-selective nitrilase-mediated transformation of 2-methyl glutaronitrile (MGN) employing *Acidovorax facilis* 72W cells (DiCosimo *et al*., 2000). When compared to existing chemical process in which MGN is directly converted via hydrogenation to a mixture of 1,3-DMPD and 1,5-DMPD, the chemoenzymatic method generates less waste and produces a single lactam isomer at significantly higher yield. BASF and Mitsubishi Rayon use nitrilase-mediated processes for the production of (*R*)-(−)-mandelic acid and its derivatives at multiton scale (Endo and Tamura, 1993; Ress-Loschke *et al*., 1998). The advantage of this process over its chemical counterpart is that no organic solvent is required and a high degree of stereoselectivity is achieved because of the presence of the biocatalyst.

3.2. Bioremediation

Synthetic nitrile compounds are widespread in the environment in the form of industrial wastewaters. Most of these are toxic, carcinogenic and mutagenic in nature (Pollak *et al*, 1991), thus there is a need to control their release into the environment. A mixed culture of bacteria containing different nitrile hydrolyzing enzymes (including NHase, amidase and nitrilase) able to metabolise effluents containing acrylonitrile, fumaronitrile, succinonitrile, *etc*. were grown in batch and continuous cultures on these components of industrial waste. A reduction in COD (75%) and significant removal (99%) of detectable toxic components was achieved by biodegradation of the effluent from acrylonitrile manufacturing industries using mixed cultures of bacteria (Wyatt and Knowles, 1995). The use of specialized consortia of micro-organisms to degrade toxic wastes therefore could be a viable alternative approach to the classical activated sludge system. Prolonged exposure to nitrile herbicides [dichlobenil (2,6-dichlorobenzonitrile), bromoxynil (3,5-dibromo-4-hydroxybenzonitrile) etc] results in symptoms of weight loss, fever, vomiting, headache and urinary problems (Freyssinet *et al*, 1996). Nitrile-metabolising enzymes efficiently degrade these cyano group-containing herbicides and prevent them from entering the food chain. *Agrobacterium radiobacter*, a bromoxynil-degrading soil bacterium, was used for the degradation of the herbicide under non-sterile batch and continuous conditions yielding a 65% reduction in the bromoxynil concentration in a column reactor after 5 days. The efficacy of degradation is enhanced by the addition of ferrous, cobaltous or cupric ions (Muller and Gabriel, 1999). A gene encoding the nitrilase has been cloned from *Klebsiella pneumoniae* ssp. *ozaenae* (McBride *et al*, 1986) and was used to raise herbicide resistant plants (Stalker *et al*, 1988). Bromoxynil resistant transgenic plants resulting from the introduction of microbial bromoxynil-specific nitrilase genes into tomato or tobacco are already approved for commercial use (Freyssinet *et al*, 1996). Similarly, other nitrile-degrading enzymes could also be potential candidates for molecular manipulation of bio-degradative systems in plant biotechnology.

4. CONCLUSIONS

The versatile biocatalytic nature and applications of nitrile converting enzymes are now increasingly recognized for the production of several pharmaceutically important compounds and fine chemicals. This relatively new biocatalytic technique also shows promise for use as a method for the preparation of *enantio-, regio-, chemo*-selective amides or acids, the synthesis of which is not feasible by chemical routes. Commercial scale processes involving kiloton syntheses of acrylamide, nicotinamide and nicotinic acid using these enzymes are excellent examples of this methodology. By virtue of their ability to eliminate highly toxic nitriles, these nitrile-degrading enzymes can also play a significant role in protecting the environment. Advances in our understanding of biosynthetic regulation, genetics and the structure

and reaction mechanism of nitrile metabolising enzymes will lead to the improved properties such as greater enzyme activity, increased tolerance to harsh compounds and improved thermostability of the biocatalysts used in commercial processes. The development of novel nitrile degrading enzymes with increasing *regio*-, *stereo*-, and *chemo*-specificities is also required for bioremediation. With continuing development in techniques related to screening, cultivation, protein and genetic engineering it is possible to isolate novel organisms of extremophilic character or improve the functions of known organisms. Though recent advances have broadened the scope of the potential applications of these versatile biocatalysts, further application-oriented studies are required to fully exploit their biotechnological potential.

REFERENCES

Almatawah Q.A., Cramp R. and Cowan D.A. (1999). Characterization of an inducible nitrilase from a thermophilic *Bacillus*. Extremophiles *3*, 283–291.

Banerjee A., Sharma R. and Banerjee U.C. (2002). Nitrile degrading enzymes: current status and future prospects. Appl. Microbiol. Biotechnol. *60*, 33–44.

Banerjee A., Sharma R. and Banerjee U.C. (2003a). A rapid and sensitive fluorometric assay method for the determination of nitrilase activity. Biotechnol. Appl. Biochem. *37*, 289–293.

Banerjee A., Kaul P., Sharma R. and Banerjee U.C. (2003b). A rapid colorimetric assay for enantioselective screening of nitrilase producing micro-organisms using pH sensitive indicators. J. Biomol. Screen. *8*, 269–275.

Bartel B. and Fink G.R. (1994). Differential regulation of an auxin producing nitrilase gene family in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA *91*, 6649–6653.

Bartling D., Seedorf M., Mithofer A. and Weiler E.W. (1992). Cloning and expression of an *Arabidopsis* nitrilase which can convert indole 3-acetonitrile to the plant hormone indole 3-acetic acid. Eur. J. Biochem. *205*, 417–424.

Baumann M., Hauer B.H. and Bornscheuer U.T. (2000). Rapid screening of hydrolases for the enantioselective conversion of difficult to resolve substrates. Tetrahedron: Asymmetry *11*, 4781–4790.

Bestwick L.A., Gronning L.M., James D.C., Bones A. and Rossiter J.T. (1993). Purification and characterization of a nitrilase from *Brassica nupus*. Physiol. Plant *89*, 811–816.

Bhalla T.C., Miura A., Wakamoto A., Ohba Y. and Furuhashi K. (1992). Asymmetric hydrolysis of alpha aminonitriles to optically active amino acids by a nitrilase of *Rhodococcus rhodochrous* PA-34. Appl. Microbiol. Biotechnol. *37*, 184–190.

Brenner C. (2002). Catalysis in the nitrilase superfamily. Curr. Opin. Strc. Biol. *12*, 775–782.

Cawston T. (1996). Metalloproteinase inhibitors and the prevention of connective tissue breakdown. Pharmacol. Ther. *70*, 163–182.

Chassin C. (1996). A biotechnological process for the production of nicotinamide. Chim. Oggi *14*, 9–12.

Chebrou H., Bigey F., Arnaud A. and Glazy P. (1996). Study of the amidase signature group. Biochem. Biophys. Acta. *1298*, 285–293.

Dadd M.R., Sharp D.C.A., Pettman A.J. and Knowles C.J. (2000). Real-time monitoring of nitrile biotransformations by mid-infrared spectroscopy. J. Microb. Meth. *41*, 69–75.

DiCosimo R., Fallon R.D., Gavagan J.E. and Herkes F.E. (2000) US Patent 6066490, to E.I. DuPont, USA.

Endo T. and Tamura K. (1993) US Patent 5223416, to Nitto Chemical Industry Co. Ltd., Japan.

Endo I., Okada M. and Yohda M. (1999). An enzyme controlled by light: the molecular mechanism of photoreactivity in nitrile hydratase. Trends in Biotechnol. *17*, 244–248.

Faber K. (1997) Biotransformation in organic chemistry: A textbook. Berlin: Springer.

Fawcett J.K. and Scott J.E. (1960). A rapid and precise method for the determination of urea. J. Clin. Pathol. *13*, 156–159.

Fournand D., Bigey F. and Arnaud A. (1998). Acyl transfer activity of an amidase from *Rhodococcus* sp. strain R312: formation of a wide range of hydroxamic acids. Appl. Environ. Microbiol. *64*, 2844–2852.

Freyssinet G., Peleissier B., Fressinet M. and Delon R. (1996). Crops resistant to oxynils: From laboratory to market. Field Crops Res. *45*, 125–133.

Gao W., Mitsuya H., Driscoll J. and Johns D. (1995). Enhancement by hydroxyurea of the antihuman immunodefficiency virus type 1 potency of 2'- ß-fluoro-2', 3'dideoxyadenosine in peripheral blood mononuclear cells. Biochem. Pharmacol. *50*, 274–276.

Gibbons B.H. and Edsall J.T. (1963). Rate of hydration of carbon dioxide and dehydration of carbonic anhydrase at 25 °C. J. Biol. Chem. *238*, 3502–3507.

Godtfredsen S.E., Ingvorsen K., Yde B. and Anderson O. (1985). Biocatalysis in organic synthesis. Amsterdam: Elsevier.

Goldhust A. and Bohak Z. (1989). Induction, purification and characterization of the nitrilase of *Fusarium oxysporum*. Biotechnol. Appl. Biochem. *11*, 581–601.

Harper D.B. (1977). Fungal degradation of aromatic nitriles: Enzymology of C-N cleavage by *Fusarium solani*. Biochem. J. *167*, 685–692.

Hashimoto Y., Kobayashi E., Endo T., Nishiyama M. and Horinouchi S. (1996). Conversion of a cyanohydrin compound into *S*-(−)-3-phenyllactic acid by enantioselective hydrolytic activity of *Pseudomonas* sp. BC-18. Biosci. Biotech. Biochem. *60*, 1273–1283.

Hillebrand H., Bartling D. and Weiler E.W. (1998). Structural analysis of the nit2/nit1/nit3 gene cluster encoding nitrilases, enzymes catalyzing the terminal activation step in indole 3-acetic acid biosynthesis in *Arabidopsis thaliana*. Plant Mol. Biol. *36*, 89–99.

Hook R.H. and Robinson W.G. (1964). Ricinine nitrilase II. Purification and properties. J. Biol. Chem. *239*, 4263–4267.

Huang W., Jia J., Cummings J., Nelson M., Schneder G. and Lindqvist Y. (1997). Crystal structure of nitrile hydratase reveals a novel iron centre in a novel fold.rystal structure of nitrile hydratase reveals a novel iron centre in a novel fold. Structure *5*, 691–699.

Kato Y., Nakamura K., Sakiyama H., Mathew S.G. and Asano Y. (2000). Novel heme-containing lyase, phenylacetaldoxime dehydratase from *Bacillus* Sp. strain OxB-1: Purification, characterization and molecular cloning of the gene. Biochemistry *39*, 800–809.

Kaul P., Banerjee A. and Banerjee U. (2004a). Biocatalysis: Applications and potential for the pharmaceutical industry. Drug Discov. World Spring Issue: 80–86.

Kaul P., Banerjee A., Mayilraj S. and Banerjee U.C. (2004b). Screening for enantioselective nitrilases: kinetic resolution of racemic mandelonitrile to (*R*)-(−)-mandelic acid by new bacterial isolates. Tetrahedron: Asymmetry *15*, 207–211.

Khachadurian A., Knox W.E. and Cullen A.M. (1960). Colorimetric ninhydrin method for total alpha amino acids of urine. J. Lab. Clin. Med. *56*, 321–325.

Kobayashi K., Yanaka N., Nagasawa T. and Yamada H. (1992). Primary structure of an aliphatic nitrilase from *Rhodococcus rhodochrous* K22 and expression of its gene and identification of its active site residue. Biochemistry *31*, 9000–9007.

Kobayashi M., Izui H., Nagasawa T. and Yamada H. (1993). Nitrilase in biosynthesis of the plant hormone indole-3-acetic acid from indole-3-acetonitrile: Cloning of the *Alcaligenes* gene and site directed mutagenesis of cysteine residues. Proc. Natl. Acad. Sci. USA *90*, 247–251.

Kobayashi M. and Shimizu S. (1994). Versatile nitrilases: Nitrile-hydrolysing enzymes. FEMS Microbiol. Lett. *120*, 217–224.

Kobayashi M., Goda M. and Shimizu S. (1998). Nitrilase catalyzes amide hydrolysis as well as nitrile hydrolysis. Biochem. Biophys. Res. Comm. *253*, 662–666.

Komeda H., Hori Y., Kobayashi M. and Shimizu S. (1996a). Transcriptional regulation of the *Rhodococcus rhodochrous* J1 nitA gene encoding a nitrilase. Proc. Natl. Acad. Sci. USA *93*, 10572–10577.

Komeda H., Kobayashi M. and Shimizu S. (1996b). A novel gene cluster including the *R. rhodochrous* J1 *nhi* BA genes encoding a new low molecular mass nitrile hydratase (L-NHase) induced by it's reaction product. J. Biol. Chem. *271*, 15796–15802.

Krallmann-Wenzel U. (1985). An improved method of ammonia determination, applicable to amidases and other ammonia producing enzyme systems of mycobacteria. Am. Rev. Respir. Dis. *131*, 432–434.

Kruse J.M. and Mellon M.J. (1953). Colorimetric determination of ammonia and cyanate. Anal. Chem. *25*, 1188–1192.

Layh N., Parratt J. and Willetts A. (1998). Characterization and partial purification of an enantioselective arylacetonitrilase from *Pseudomonas fluorescens* DSM 7155. J. Mol. Cat. B: Enz. *5*, 467–474.

Lebuhn M. and Hartmann A. (1993). Method for the determination of indole-3-acetic acid and related compounds of L-tryptophan catabolism in soils. J. Chromat. A *629*, 255–266.

Lowry O.H., Roberts N.R., Wu M.L., Hixon W.S. and Crawford E.J. (1954). The quantitave histochemistry of brain II - Enzymatic measurements. J. Biol. Chem. *207*, 19–37.

Martin R.G. and Rosner J.L. (2001). The AraC transcrptional activators. Curr. Opin. Microbiol. *4*, 132–137.

McBride K.E., Kenny J.W. and Stalker D.M. (1986). Metabolism of the herbicide bromoxynil by *Klebsiella pneumomiae* subsp. *ozanae*. Appl. Environ. Microbiol. *52*, 325–330.

Muller D. and Gabriel J. (1999). Bacterial degradation of the herbicide bromoxynil by *Agrobacterium radiobacter* in biofilm. Folia Mikrobiologia *44*, 377–379.

Nagasawa T., Takeuchi K., Nardi-Devi V., Mihara Y. and Yamada H. (1991). Optimum culture conditions for the production of cobalt containing nitrile hydratase by *Rhodococcus rhodochrous* J1. Appl. Microbiol. Biotechnol *34*, 783–788.

Nagasawa T. (1993). The superiority of the third generation catalyst, *Rhodococcus rhodochrous* J1 nitrile hydratase, for industrial production of acrylamide. Appl. Microbiol. Biotechnol. *40*, 189–195.

Nagasawa T., Wieser M., Nakamura T., Iwahara H., Yoshida T. and Gekko K. (2000). Nitrilase of *Rhodococcus rhodochrous* J1, conversion into the active form by subunit association. Eur. J. Biochem. *267*, 138–144.

Nakai T., Hasegawa T., Yamashita E., Yamamoto M., Kumaska T., Ueki T., Nanba H., Ikenaka Y., Sato M. and Tsukihara T. (2000). Crystal structue of N-carbamyl-D-amino acid amidohydrolase with a novel catalytic framework common to amido hydrolases. Structure *8*, 729–739.

Nawaz M., Khan A., Bhattacharya D., Sittoneu P. and Cerniglia C. (1996). Physical, biochemical, and immunological characterization of a thermostable amidase from *Klebsiella pneumoniae* NCTR 1. J. Bacteriol. *178*, 2397–2401.

Nishise H., Kurihara M. and Tani Y. (1987). Microbial synthesis of tranexamic acid intermediate from dinitrile. Agric. Biol. Chem. *51*, 2613–2616.

O'Reilly C. and Turner P.D. (2003). The nitrilase family of CN hydrolysing enzymes – a comparative study. J. Appl. Microbiol. *95*, 1161–1174.

Pace H.C. and Brenner C. (2001). The nitrilase superfamily: Classification, structure and function. Genome Biol. *2*, 1–9.

Patel R.N. (2001). Biocatalytic synthesis of intermediates for the synthesis of chiral drug substances. Curr. Opin. Biotechnol. *12*, 587–604.

Payne M., Wu S., Fallon R., Tudor G., Turner I., Steglitz B. and Nelson M. (1997). A stereoselective cobalt containing nitrile hydratase. Biochemistry *36*, 5447–5454.

Piotrowski M., Schonfelder S. and Weiler E.W. (2001). The *Arabidopsis thaliana* isogene NIT4 and its orthologs in tobacco encode beta cyano L-alanine hydratase/nitrilase. J. Biol. Chem. *276*, 2616–2621.

Pollak P., Romender G., Hagedorn F. and Gelbke H.P. (1991) Ullmans encyclopaedia of industrial chemistry. Weinheim: Wiley – VCH.

Ress-Loschke M., Friedrich T., Hauer B., Mattes R. and Engels D. (1998) DE198448129A1, to BASF.

Schill S. Levy, Soubrier F., Coq A.M.C.-L., Faucher D., Crouzet J. and Petre D. (1995). Aliphatic nitrilase from a soil isolated Comamonas testoteroni sp. gene cloning and overexpression, purification and primary structure. Gene *161*, 15–20.

Schmid A., Dordick J.S., Hauer B., Kiener A., Wubbolts M. and Witholt B. (2001). Industrial biocatalysis today and tomorrow. Nature *409*, 258–268.

Simons S.S. and Johnson D.F. (1976). The structure of the fluorescent adduct formed in the reaction of o-phthalaldehyde and thiols with amines. J. Am. Chem. Soc. *98*, 7098–7099.

Simons S.S. and Johnson D.F. (1978). Reaction of o-phthalaldehyde and thiols with primary amines: formation of 1-alkyl(and aryl)thio-2-alkylisoindoles. J. Org. Chem. *43*, 2886–2891.

Solomonson L.P. and Spehar A.M. (1981) Cyanide as a metabolic inhibitor. London: Academic Press.

Stalker D.M., Mcbride K.E. and Malyj L.D. (1988). Herbicide resistence in transgenic plants expressing a bacterial detoxification gene. Science *242*, 419–423

Sugai T., Yamazaki T., Yokoyama M. and Ohta H. (1997). Biocatalysis in organic synthesis: The use of nitrile and amide-hydrolyzing micro-organisms. Biosci. Biotech. Biochem. *61*, 1419–1427.

Sugiura Y., Kuwahara J., Nagasawa T. and Yamada H. (1987). Nitrile hydratase: the first non-heme iron enzyme with a typical low spin Fe(III) active centre. J. Am. Chem. Soc. *109*, 5848–5850.

Thimann K.V. and Mahadevan S. (1964). Nitrilase, its substrate specificity and possible mode of action. Arch. Biochem. Biophys. *107*, 62–68.

Tsunoda H. and Yamaguchi K. (1995). The cDNA sequence of an auxin producing nitrilase homologue in tobacco. Plant Physiol. *109*, 339.

Wu Z.L. and Li Z.Y. (2003). Biocatalytic asymmetric hydrolysis of (±)-beta hydroxy nitriles by *Rhodococcus* sp. CGMCC 0497. J. Mol. Cat. B: Enz. *22*, 105–112.

Wyatt J.M. and Knowles C.J. (1995). Microbial degradation of acrylonitrile waste effluents: the degradation of effluents and and codendsates from the manufacture of acrylonitrile. Int. Biodeter. Biodegrad. 227–248.

Yamada H. and Nagasawa T. (1994) US Patent 5334519, to Nitto Chemical Industry Co. Ltd. Japan.

Yamada H. and Kobayashi M. (1996). Nitrile hydratase and it's application to industrial production of acrylamide. Biosci. Biotech. Biochem. *60*, 1391–1400.

Yamamoto K., Ueno Y., Otsubo K., Kawakami K. and Komatsu K. (1990). Production of (*S*)-(+)-ibuprofen from a nitrile compound by *Acinetobacter* sp. AK 226. Appl. Environ. Microbiol. *56*, 3125–3129.

Yamamoto K., Ueno Y., Otsubo K., Yamane H., Komatsu K. and Tani Y. (1992). Efficient conversion of dinitrile to mononitrile -monocarboxylic acid by *Corynebacterium* sp. C5 cells during tranexamic acid synthesis. J. Ferment. Bioeng. *73*, 125–129.

CHAPTER 31

ASPARTASES: MOLECULAR STRUCTURE, BIOCHEMICAL FUNCTION AND BIOTECHNOLOGICAL APPLICATIONS

TOMOHIRO MIZOBATA AND YASUSHI KAWATA*

Department of Biotechnology, Faculty of Engineering, and Department of Biomedical Science, Institute of Regenerative Medicine and Biofunction, Graduate School of Medical Science, Tottori University, Koyama-Minami, Tottori, Japan

**kawata@lemon.bio.tottori-u.ac.jp*

1. INTRODUCTION

The experimental finding that a unique enzyme in bacteria mediates an equilibrium between L-aspartic acid, fumaric acid and ammonia was first reported by Quastel and Woolf (Quastel and Woolf, 1926) using cultures of *E. coli* shifted to resting phase by the addition of propyl alcohol or toluene. Further characterization determined that this enzyme was capable of catalysing the deamination of L-aspartate, and it was named aspartase in 1929 (Cook and Woolf, 1928). Its isolation was successfully achieved from cell extracts of *Pseudomonas fluorescens* (Virtanen and Tarnanen, 1932), and was subsequently followed by a series of studies performed by Ellfolk describing the initial characterization of this activity which involved its partial purification (Ellfolk, 1953a) and the determination of its substrate specificity (Ellfolk, 1954) and chemical modification (Ellfolk, 1953b).

The purification of aspartase from *E. coli* cells was first reported by Rudolph and Fromm (Rudolph and Fromm, 1971) and involved multiple purification steps. Subsequent refinements in the protocol gradually succeeded in improving both the yield and purity of the enzyme, and an important addition to the purification protocol, dye-ligand chromatography using Red-A agarose and elution with 1 mM L-aspartate (Karsten *et al*, 1985), allowed researchers to obtain pure aspartase in very high yields. The cloning of the structural gene of this enzyme was successful for a variety of bacterial species including *E. coli* (Guest *et al*, 1984) and *Pseudomonas*

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 549–565

```
                1          11         21         31         41         51
E. coli         M------SNN IRIEEDLLGT REVPADAYYG VHTLRAIENF YISNNKISDI PEFVRGMVMV
P. fluorescens  MISVMSSAAS FRTEKDLLGV LEVPAQAYYG IQTLRAVNNF RLSGVPISHY PKLVVGLAMV
Cytophaga KUC-1 M-------GS TRKEHDFLGE LDIPNHLYYG IQTFRAVENF NITGIPISKE PLFIKALGYV
B. subtilis     M---LNGQKE YRVEKDFLGE KQIEADVYYG IQTLRASENF PITGYKI--H EEMINALAIV
Bacillus YM55-1 M------NTD VRIEKDFLGE KEIPKDAYYG VQTIRATENF PITGYRI--H PELIKSLGIV
                *        .  * *.*:**   ::  . *** ::*:** :**  ::.  *      :: .:  *

                61         71         81         91         101        111
E. coli         KKAAAMANKE LQTIPKSVAN AIIAACDEVL NNGKCMDQFP VDVYQGGAGT SVNMNTNEVL
P. fluorescens  KQAAADANRE LGQLSERKHA AISEACARLI -RGDFHEEFV VDMIQGGAGT STNMNANEVI
Cytophaga KUC-1 KKAAALANKD CGRLDPKIAE AICYGSDQVI -AGKFDQEFV SDLIQGGAGT SVNMNANEVI
B. subtilis     KKAAALANMD VKRLYEGIGQ AIVQAADEIL -EGKWHDQFI VDPIQGGAGT SMNMNANEVI
Bacillus YM55-1 KKSAALANME VGLLDKEVGQ YIVKAADEVI -EGKWNDQFI VDPIQGGAGT SINMNANEVI
                *:.** ** :    :        *  .. .::   *.  ::*   *  ****** * ***:***:

                121        131        141        151        161        171
E. coli         ANIGLELMGH QKGEYQYLNP NDHVNKCQST NDAYPTGFRI AVYSSLIKLV DAINQLREGF
P. fluorescens  ANIALEAMGH QKGEYQYLHP NNDVNMAQST NDAYPTAIRL GLLLGHDALL ASLDSLIQAF
Cytophaga KUC-1 ANIGLEYLGH KKGDYNFLHP NNHVNCSQST NDAYPSAFRI ALYLKMESFI KTLEGLEVAF
B. subtilis     GNRALEIMGH KKGDYIHLSP NTHVNMSQSQ NDVFPTAIHI STLKLLEKLL KTMEDMHSVF
Bacillus YM55-1 ANRALELMGE EKGNYSKISP NSHVNMSQST NDAFPTATHI AVLSLLNQLI ETTKYMQQEF
                .* .** :*. :**:*  : * * .** .**  **.:*:. :: .       ::  : . :   *

                181        191        201        211        221        231
E. coli         ERKAVEFQDI LKMGRTQLQD AVPMTLGQEF RAFSILLKEE VKNIQR-TAE LLLEVNLGAT
P. fluorescens  AAKGAEFSHV LKMGRTQLQD AVPMTLGQEF RAFATTLGED LARLKTLAPE LLTEVNLGGT
Cytophaga KUC-1 VANGEEFKSV LKMGRTQLQD AVPMTLGQEF RSYATTIGED VRRLKE-AQS LVLEINMGAT
B. subtilis     KQKAQEFHSV IKMGRTHLQD AVPIRLGQEF EAYSRVLERD IKRIKQ-SRQ HLYEVNMGAT
Bacillus YM55-1 MKKADEFAGV IKMGRTHLQD AVPILLGQEF EAYARVIARD IERIAN-TRN NLYDINMGAT
                  :. **  : :*****:*** ***: ***** .:::  : .: : .:   : .  : ::*:*.*

                241        251        261        271        281        291
E. coli         AIGTGLNTPK EYSPLAVKKL AEVTGFPCVP AEDLIEATSD CGAYVMVHGA LKRLAVKMSK
P. fluorescens  AIGTGINADP RYQALAVQRL ATISGQPLVP AADLIEATSD MGAFVLFSGM LKRTAVKLSK
Cytophaga KUC-1 AIGTRVNAPE GYPEICVNYL AKEVGIPLTL SPDLIEATVD TGAYVQIMGT LKRTAVKISK
B. subtilis     AVGTGLNADP EYIKQVVKHL ADISGLPLVG ADHLVDATQN TDAYTEVSAS LKVCMMNMSK
Bacillus YM55-1 AVGTGLNADP EYISIVTEHL AKFSGHPLRS AQHLVDATQN TDCYTEVSSA LKVCMINMSK
                *:** :*:    *    .: * *   * *    : .*::** :  ..:. . .  **   :::**

               301        311        321        331        341        351
E. coli         ICNDLRLLSS GPRAGLNEIN LPELQAGSSI MPAKVNPVVP EVVNQVCFKV IGNDTTVTMA
P. fluorescens  ICNDLRLLSS GPRTGINEIN LPARQPGSSI MPGKVNPVIP EAVNQVAFQV IGNDLALTMA
Cytophaga KUC-1 ICNDLRLLSS GPRTGFNEIN LPARQPGSSI MPGKVNPVIP EVVNQTCFYV IGQDLTVTMA
B. subtilis     IANDLRLMAS GPRAGLAEIS LPARQPGSSI MPGKVNPVMA ELINQIAFQV IGNDNTICLA
Bacillus YM55-1 IANDLRLMAS GPRAGLSEIV LPARQPGSSI MPGKVNPVMP EVMNQVAFQV FGNDLTITSA
                *.*****::* ***:*: **  **  *.**** **.*****:. * :** .* * :*:* ::  *

                361        371        381        391        401        411
E. coli         AEAGQLQLNV MEPVIGQAMF ESVHILTNAC YNLLEKCING ITANKEVCEG YVYNSIGIVT
P. fluorescens  AEGGQLQLNV MEPLIAFKIF DSIRLLQRAM DMLREHCIVG ITANEARCRE LVEHSIGLVT
Cytophaga KUC-1 AEAGQLQLNV MEPVIAFAMF TSLDYLSNAI QTLIDKCIIG ITANVDHCYN MVMNSIGIVT
gi|142518|gb|AA SEAGQLELNV MEPVLVFNLL QSISIMNNGF RSFTDNCLKG IEANEKRMKQ YVEKSAGVIT
Bacillus YM55-1 SEAGQFELNV MEPVLFFNLI QSISIMTNVF KSFTENCLKG IKANEERMKE YVEKSIGIIT
                :*.**::*** ***::   ::  *:  : .     : ::*: * * **        * :* *::*

                421        431        441        451        461        471
E. coli         YLNPFIGHHN GDIVGKICAE TGKSVREV-V LERGLLTEAE LDDIFSVQNL MHPAYKAKRY
P. fluorescens  ALNPYIGYEN ATRIARIALE SGRGVLEL-V REEGLLDDAM LDDILRPENM IAPRLVPLKA
Cytophaga KUC-1 QLNPILGYEI SASIAGEALK MNKSVHEIVV VERKLITQEK WDEIYSLDNL INPKFITK--
B. subtilis     AVNPHLGYEA AARIAREAIM TGQSVRDL-C LQHDVLTEEE LDIILNPYEM TKPGIAGKEL
Bacillus YM55-1 AINPHVGYET AAKLAREAYL TGESIREL-C IKYGVLTEEQ LNEILNPYEM THPGIAGRK-
                 :** :*:.  .  :.  .    ...: ::    :  :: :    : *    ::   *

                481
E. coli         TDESEQ
P. fluorescens  ------
Cytophaga KUC-1 ------
B. subtilis     LEK---
Bacillus YM55-1 ------
```

fluorescens (Takagi *et al*., 1986), and overproduction of the gene product in suitable hosts followed soon after. At present, researchers are able to obtain highly purified samples of aspartase from a variety of organisms using the combination of overproducing strains and dye-ligand chromatography.

A comparatively recent development has been the isolation and characterization of aspartases from a number of extremophiles, for example the moderately thermophilic *Bacillus* species *Bacillus* sp. YM55-1 (Kawata *et al*., 1999) and the psychrophile *Cytophaga* sp. (Kazuoka *et al*., 2003). The aspartases isolated from these exotic sources possess numerous unique characteristics compared to the aspartases from mesophilic sources. Details will be given in the following sections.

2. BASIC CHARACTERISTICS OF ASPARTASES

Aspartases from different organisms share a number of common structural characteristics. The enzymes from *Pseudomonas fluorescens* (Takagi *et al*., 1986), *E. coli* (Takagi *et al*., 1985), *Bacillus* sp. YM55-1 (Kawata *et al*., 1999) and *Cytophaga* (Kazuoka *et al*., 2003) are all homotetrameric enzymes with subunit molecular weights of around 50∼52 kDa. Amino acid sequence homologies are extensive (Fig. 1), and 45% ∼ 50% identity occurs between sequences from various sources (Sun and Setlow, 1991; Kawata *et al*., 2000). The *Bacillus subtilis* and *Bacillus* sp. YM55-1 aspartases exhibit 71% identity (Kawata *et al*., 2000). In addition, the aspartases are members of a larger family of enzymes which includes the type II fumarases (Acuna *et al*., 1991) and the adenylosuccinate lyases (Woods *et al*., 1988). High levels of amino acid sequence homology are also shared between these enzymes.

As determined from gene sequencing the *E. coli* aspartase subunit is composed of 478 amino acid residues and has a molecular weight of approximately 52 kDa (Takagi *et al*., 1985). The theoretical pI of this enzyme is 5.2, and a T_m of 55 °C was determined in thermal unfolding experiments monitored by intrinsic tyrosine fluorescence (Kawata *et al*., 1999). The enzyme retains 50% of its initial activity after a 4 minute incubation at this temperature but is completely inactivated after 30 min (Kawata *et al*., 1999).

Figure 1. Alignment of the amino acid sequences of aspartases. The numbering scheme shown is arbitrary and does not correspond to the residue numbers of any specific type of aspartase. Alignment of five known amino acid sequences are shown: E. coli, *Escherichia coli* (Burland *et al*., 1995); P. fluorescens, *Pseudomonas fluorescens* (Takagi *et al*., 1986); Cytophaga KUC-1, *Cytophaga* sp. KUC-1 (Kazuoka *et al*., 2003); B. subtilis, *Bacillus subtilis* (Sun and Setlow, 1991); and Bacillus YM55-1, *Bacillus* sp. YM55-1 (Kawata *et al*., 2000). In the case of *Escherichia coli* aspartase the first 15 residues of the derived amino acid sequence were omitted to reflect the actual amino acid sequence of the aspartase protein. Alignment analysis was performed using an online version of the MAFFT sequence alignment program (Katoh *et al*., 2005)

2.1. Enzymatic Characteristics

The enzymatic characteristics of the *E. coli* aspartase are very notable in many respects. Firstly, the specificity of this enzyme for its substrates, L-aspartate and fumarate, is extremely high; to date, only the suicide substrate L-aspartate β-semialdehyde has been discovered that is efficiently deaminated by this enzyme (Yumoto *et al*., 1982; Schindler and Viola, 1994). Aspartase is inhibited quite efficiently however, by a variety of compounds that mimic the structural properties of L-aspartate, such as D- and L-malate, succinate, and *N*-acetyl L-aspartate. Falzone and co-workers have proposed that the presence of a carboxylate or equivalent functional group in the 1 and 4 positions of the backbone carbon chain of the substrate is important for binding to the aspartase binding site (Falzone *et al*., 1988).

Another interesting characteristic of this enzyme is its behaviour at various pH and the effects of different divalent cations. In the absence of a divalent cation the pH dependence of aspartase displays a bell-shaped curve with decreased activity at the acidic and alkaline extremes and a single activity maximum at pH $\sim$ 6 (Karsten and Viola, 1991). However, in the presence of 10 mM Mg^{2+}, the activity at alkaline pH is greatly increased and a second activity maximum is observed at pH $\sim$ 8. The presence of divalent cations acts as an activator of aspartase activity at high pH. Further studies revealed the nature of this cation-derived activation to be a relatively complicated phenomenon. In this pH region, the substrate saturation curves display a distinctly sigmoidal character, indicative of a cooperative mechanism involving multiple subunits of the aspartase tetramer (Ida and Tokushige, 1985; Karsten *et al*., 1986). When this characteristic was examined it was found that L-aspartate acts in cooperation with divalent cations to allosterically enhance the activity of aspartase (in both directions of the reversible deamination-amination reaction). Thus L-aspartate is not only a substrate of aspartase, but in cooperation with divalent cations it is also a potent activator. Similar behaviour has been confirmed for the related protein fumarase C, for which very high concentrations of the substrate malate result in enhanced activity (Beeckmans and Van Driessche, 1998; Rose and Weaver, 2004). pH dependent enzyme function, divalent cation action, and the presence of a secondary substrate binding site which enhances the activity of both the forward and backward reactions, all combine to confer aspartase an exquisitely complex catalytic mechanism.

The enzymatic mechanism of aspartase was elucidated by studies which involved the characterization of a specific inhibitor of aspartase and the analysis of isotope effects using specifically labelled substrate molecules. In 1980 Porter and Bright identified 2-nitro-3-aminopropionate to be a very potent inhibitor of the aspartate deamination activity of aspartase (Porter and Bright, 1980). This compound was found to inhibit aspartase most strongly under pH conditions where the *aci*-acid form was most stable. This *aci*-acid form closely resembles a carbanion intermediate state of L-aspartate with a hydrogen atom removed, and this was taken as strong evidence for the formation of a carbanion intermediate during the reaction cycle. Detailed studies were performed on the kinetic mechanism of the aspartase

from *Hafina alvei* using deuterium-substituted and ^{15}N-substituted substrates (Nuiry *et al.*, 1984). This study determined that the kinetic mechanism is initiated by the binding of a divalent cation followed by binding of the activator aspartate molecule. Binding of cation and aspartate occur separately, rather than through a pre-formed cation-aspartate complex. A significant isotope effect attributable to ^{15}N substitution was also observed which suggested that the rate limiting step of the deamination reaction was the scission of the C-N bond of aspartate. The effects of this inhibitor, taken together with additional isotope effects observed with specifically deuterated forms of L-aspartate, indicate the following enzyme mechanism for aspartase (Fig. 2 Yoon *et al*., 1995).

The deamination of L-aspartate is initiated by the removal of a proton from C3 of L-aspartate. The subsequent intermediate state is stabilized in the carbanion form, as suggested by the effects of 2-nitro-3-aminoprionate. The next, rate limiting, step of the overall reaction is the removal of the ammonium group NH_3. The detached NH_3 molecule is then protonated by a general acid group located in the vicinity to form NH_4^+, which is the actual leaving group. The release of the products fumarate, NH_4^+ and the divalent cation are thought to occur in a random manner. This acid/base mechanism of catalysis requires three properties associated with the enzyme structure for successful catalysis to occur (Fig. 2): a general base that mediates the initial removal of the proton from C3; a residue that stabilizes the carbanion intermediate state, most likely through charge stabilization; and a general

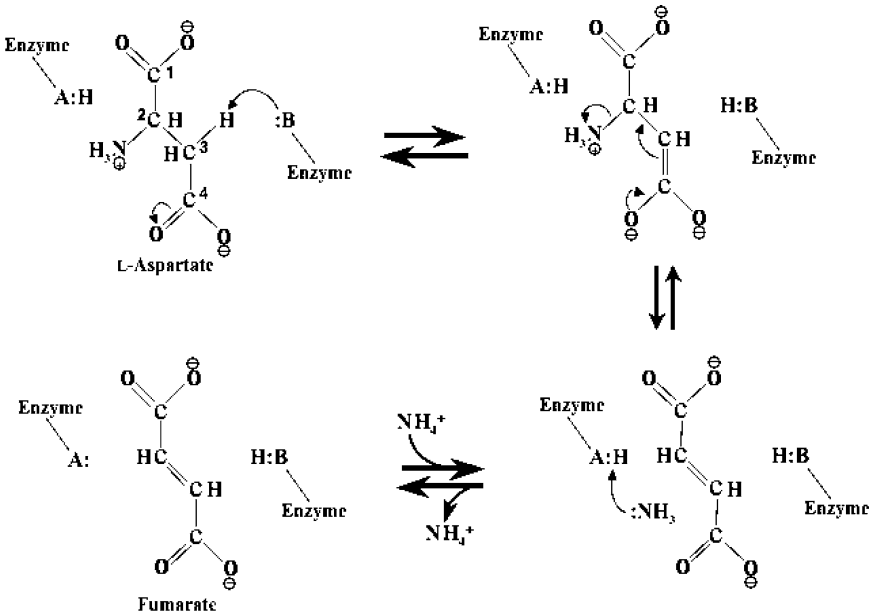


Figure 2. Proposed catalytic mechanism of aspartase (Yoon *et al*., 1995). The mechanism is depicted clockwise from the upper left-hand corner

acid group that assists the removal of the ammonia group. These three factors are considered to be assigned to specific residues located in the vicinity of the aspartase active site. The identification of these residues proved to be exceedingly difficult, however. Initial studies that utilized chemical modification to target specific amino acid types as well as detailed analysis of the pH dependence of aspartase catalysis suggested that cysteine (Mizuta and Tokushige, 1975; Ida and Tokushige, 1985), lysine (Karsten and Viola, 1991) and histidine (Ida and Tokushige, 1984) residues were important participants in the catalytic mechanism. Armed with this information, a number of groups performed site-directed mutagenesis studies on these three residues, in some cases focusing on those that were highly conserved among various species, and in the case of cysteine (Chen *et al*, 1996) an exhaustive substitution analysis of all of the relevant residues in the amino acid sequence. The results from these studies indicated that none of the cysteine residues were essential for catalysis. Similar studies also ruled out a conserved histidine residue as an active participant in catalysis (Viola, 2000). Finally, two highly conserved lysine residues, Lys 327 and Lys 55, were found (through replacement by arginine residues) to be very important for enzyme function (Saribas *et al*, 1994). Subsequently, the loss of activity in the Lys55Arg mutant was found to be due to aggregation and inactivation of the aspartase that could be partially reversed by incubation in denaturant (Jayasekera and Viola, 1999). However, Lys 327 was determined to be important in the initial binding of substrate and also in communications between the active site and the pH-dependent activation site.

2.2. X-ray Crystal Structure of Aspartase

In 1997 a high resolution crystal structure of *E. coli* aspartase, shown in Fig. 3, was reported (Shi *et al*, 1997). This was the second structure of a member of the aspartase-fumarase superfamily to be elucidated, the first being that of *E. coli* fumarase (Weaver *et al*, 1995). The aspartase subunit consists of three distinct sequentially arranged domains, denoted 1, 2 and 3 (Fig. 3a). The central domain, domain 2, consists almost entirely of α-helical elements and forms the interface for subunit association. Domains 1 and 3 display a more dynamic nature in the crystallography analyses and are thought to form the actual active centres of the enzyme. The final arrangement of the aspartase tetramer may be described as consisting of a central subunit interface region and four active centres arranged regularly on both sides of this central interface (Fig. 3b).

Three of the four subunits in aspartase are thought to contribute important amino acid residues to form a given active site (Shi *et al*, 1997). However, the actual position of the active site was not readily apparent in the initial structural determination of the apo form of aspartase. Data from the mutagenesis studies outlined above that suggested that Lys 327 was important for effective binding of substrate (Saribas *et al*, 1994) localized the active site in the vicinity of this lysine residue. A number of amino acid residues found within a 15 Å radius of Lys 327 were listed as candidates for catalytic participation (Jayasekera *et al*, 1997). Of five residues identified in this study, one, Arg 29, has been implicated in substrate binding, and

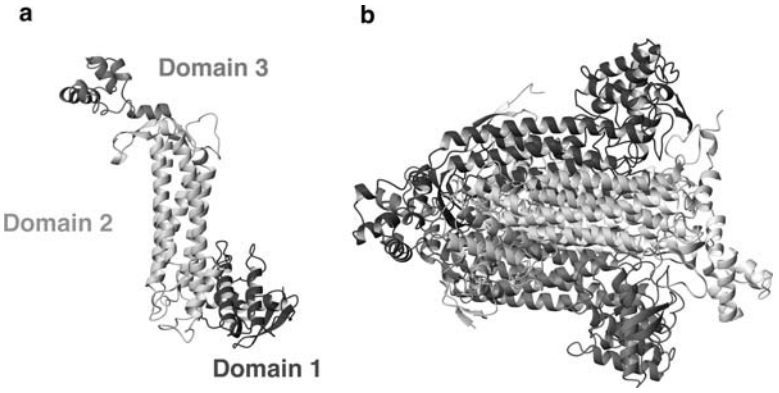


Figure 3. X-ray crystal structure of *Escherichia coli* aspartase (Shi *et al*., 1997). (a) Domain composition of the *E. coli* aspartase subunit. (b) Structure of the aspartase tetramer

another, Ser 143, has been implicated to act as a general acid catalyst that mediates the protonation of the departing ammonium group. Still missing from this molecular view of the aspartase active site is the residue that acts as a general base to remove a proton from C3 of L-aspartate to form the carbanion intermediate. Although a number of candidates have been probed to date, notably His 26 and Asp 10, neither of these residues satisfied the criteria of a general base catalyst.

3. *BACILLUS* sp. YM55-1 ASPARTASE

Relatively recently a new member of the aspartase family was isolated from an aerobic, moderately thermophilic bacterium, *Bacillus* sp. YM55-1 (Kawata *et al*., 1999). This aspartase displayed a number of characteristics that were different from those of the mesophilic aspartases. Firstly, the specific activity of *Bacillus* sp. YM55-1 aspartase was significantly (∼5 fold) higher than that of the other mesophilic aspartases under identical conditions (assayed at 30 °C). This enhancement was much more pronounced when the enzyme was assayed at its optimum temperature of 55 °C, yielding a specific activity of 2200 units/mg protein, which is among the highest specific activities known for an aspartase. The stability of this aspartase against thermal and chemical denaturation was also significantly better than that of *E. coli* aspartase. This has been attributed to increased numbers of intersubunit hydrogen bonds and ion pairs in the *Bacillus* sp. YM55-1 enzyme (Fujii *et al*., 2003). Another interesting difference between the *Bacillus* sp. YM55-1 aspartase and other mesophilic aspartases is that the former does not display enhanced activity at alkaline pH in the presence of divalent cations, nor does it undergo substrate activation by L-aspartate. The enzyme kinetics of *Bacillus* sp. YM55-1 aspartase are hyperbolic at all pH values tested, and cooperativity is notably lacking.

Cloning and overproduction of *Bacillus* sp. YM55-1 aspartase in *E. coli* (Kawata *et al*., 2000) and crystallization/determination of its structure (Fujii *et al*., 2003)

followed soon after (Fig. 4). The general structure of *Bacillus* sp. TM55-1 aspartase is essentially identical to that of *E. coli* aspartase and *E. coli* fumarase: a central subunit interface consisting mainly of α-helices, and active sites flanking this central structure (compare Figs. 3 and 4). Upon closer examination a number of additional similarities, and some notable differences, are apparent between these three molecular structures.

A docking simulation of aspartate bound to *Bacillus* sp. YM55-1 aspartase was performed in order to ascertain the orientation and possible active site residues in this enzyme (Figs. 5 and 6; Fujii *et al*., 2003). Lys 327, the residue that binds the α-carboxyl group in *E. coli* aspartase, corresponds to Lys 324 in the *Bacillus* sp. structure, and this residue was postulated to act in a similar role in this thermostable enzyme. The amino group was arranged to be bound through three separate residues (Fig. 6): Asn142 (*E. coli* Asn 145), His 188 (*E. coli* Gln 191) and either Gly 98 (*E. coli* Gly 100) or Thr 101 (*E. coli* Thr 104). Docking was based on the finding that residues homologous to these three in *E. coli* fumarase (Ser 98, Thr 100, Asn 141 and His 188) were found complexed with a water molecule. This water molecule is regarded as being very important in the catalytic mechanism of fumarase, and an analogy may be derived that implicates these residues in aspartase catalysis as well. The β-carboxyl group of docked L-aspartate was found to interact with Ser 140 (*E. coli* Ser 143) in the completed docking model (Fig. 6). This residue is located at the N-terminal end of a helix dipole (helix 6 in the subunit structure) and the positive charge provided by this dipole is thought to assist in binding.

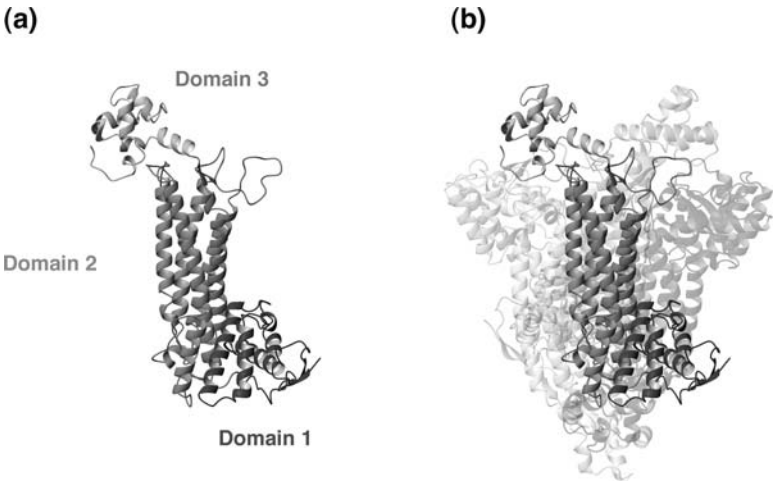


Figure 4. X-ray crystal structure of *Bacillus* sp. YM55-1 aspartase (Fujii *et al*., 2003). (a) The domain composition of the aspartase subunit, oriented in a similar manner to that of the *E. coli* aspartase subunit shown in Fig. 3a. (b) Structure of the *Bacillus* sp. YM55-1 aspartase tetramer. In this figure the tetramer was arranged in a manner such that the relative positions of the domains in a given subunit is apparent. Images were created using a coordinate file of the *Bacillus* aspartase tetramer (provided generously by Drs. T. Fujii and Y. Hata, Kyoto University)

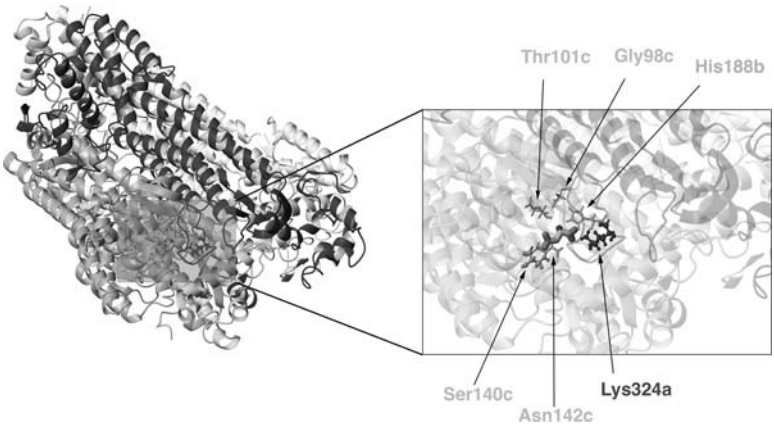


Figure 5. Close-up view of the active site of *Bacillus* sp. YM55-1 aspartase with an L-aspartate molecule in the postulated binding site. The left image shows a view of the aspartase tetramer with different shading assigned to individual subunits. Following the naming convention of Fujii *et al.* (2003) individual subunits are shaded dark to light in the following sequence: subunit a, subunit b, subunit c. Subunit d, which does not participate in the formation of the active site, is the most lightly shaded. The right image shows a zoomed view of the portion of the left panel. The relative orientation of the two images is identical. The amino acids that are postulated to participate in the catalytic mechanism of aspartase are drawn in the zoomed view in either stick or ball-and-stick form, and the relative shading of the amino acid residues denotes the subunit from which each residue is derived. The docked L-aspartate residue is shown in thick stick form

From numerous studies of members of the aspartase-fumarase superfamily it has been suggested that residues at the position corresponding to His 188 in *Bacillus* sp. YM55-1 aspartase – modeled to interact with the amino group of L-aspartate in

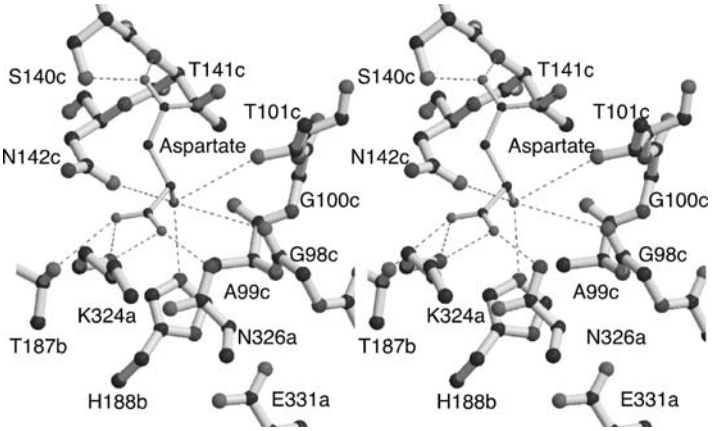


Figure 6. Stereo view of the active site of *Bacillus* sp. YM55-1 aspartase, reprinted from Fujii *et al.* (2003)

Figs. 5 and 6 – not only participate in substrate binding but also actively participate in the catalytic mechanism as a general acid catalyst (Blanchard and Cleland, 1980; Weaver and Banaszak, 1996; Weaver *et al*, 1997; Wu *et al*, 1998; Lee *et al*, 1999). His 188 is highly conserved among most of the members of the fumarase superfamily, including the aspartase from *Bacillus* sp. YM55-1. However, that is not the case in *E. coli* aspartase where a glutamine residue occurs at this position (Gln 191 in *E. coli*). Although this difference in a crucial residue may indicate that this position in the active site of the aspartase-fumarase superfamily is not an essential functional position, the homology between *Bacillus* sp. YM55-1 aspartase and the other members of the fumarase family may warrant a more detailed scrutiny of this interesting residue.

Another residue that was proposed to be an active participant in the catalytic mechanism of the *E. coli* aspartase was Ser 143 (Ser 140 in *Bacillus* sp. YM55-1 aspartase). Ser 143 was postulated to be the general acid residue that protonates the ammonia leaving group (Jayasekera *et al*, 1997). The position of Ser 140 in *Bacillus* sp. YM55-1 aspartase is, however, comparatively distant from the candidate active site region and the role of general acid catalyst seemed difficult to achieve in this enzyme (Fujii *et al*, 2003). Instead, Ser 140 is postulated to bind the β-carboxyl group of L-aspartate, thus filling an important role in substrate binding for this residue (Fig. 6). An alternate residue that could assume the role of general acid catalyst in place of Ser 140 was not readily apparent in the *Bacillus* sp. YM55-1 structure. The detailed understanding of the catalytic mechanism of *Bacillus* sp. aspartase is therefore far from complete.

With regard to the substrate activation effect that is seen in both *E. coli* aspartase and fumarase but not in *Bacillus* sp. YM55-1 aspartase, structural comparison has revealed that a site corresponding closely to the B-site in *E. coli* fumarase (where a L-malate molecule was found to be bound by crystal structure analysis (Weaver and Banaszak, 1996)) also exists in the *Bacillus* sp. YM55-1 and *E. coli* aspartases (Fig. 7). A detailed comparison of this region (corresponding to residues 123–128 in *Bacillus* sp. aspartase) showed that the structure of *Bacillus* sp. YM55-1 aspartase more closely resembles the structure of *E. coli* fumarase C (Fig. 7a) rather than the apo structure of *E. coli* aspartase (Fig. 7b). It may be that the *Bacillus* sp. aspartase is locked in an active conformation similar to that seen for ligand-bound *E. coli* fumarase. An additional striking finding was that in *Bacillus* sp. YM55-1 aspartase, a tyrosine side chain (Tyr 126) was found in a position which partially occluded this activation site (Fig. 7a). This may be the reason why the enzyme does not display a prominent substrate activation effect and is locked in a permanently active conformation. Related to this finding, Fujii and co-workers showed that the structure of apo *E. coli* aspartase differs markedly from the structures of *Bacillus* sp. YM55-1 aspartase and *E. coli* fumarase (liganded form) in the regions corresponding to residues 40–45, 80–84, 228–241 and 265–272. A detailed examination of these regions may yield additional clues about substrate-activated catalysis in *E. coli* aspartase.

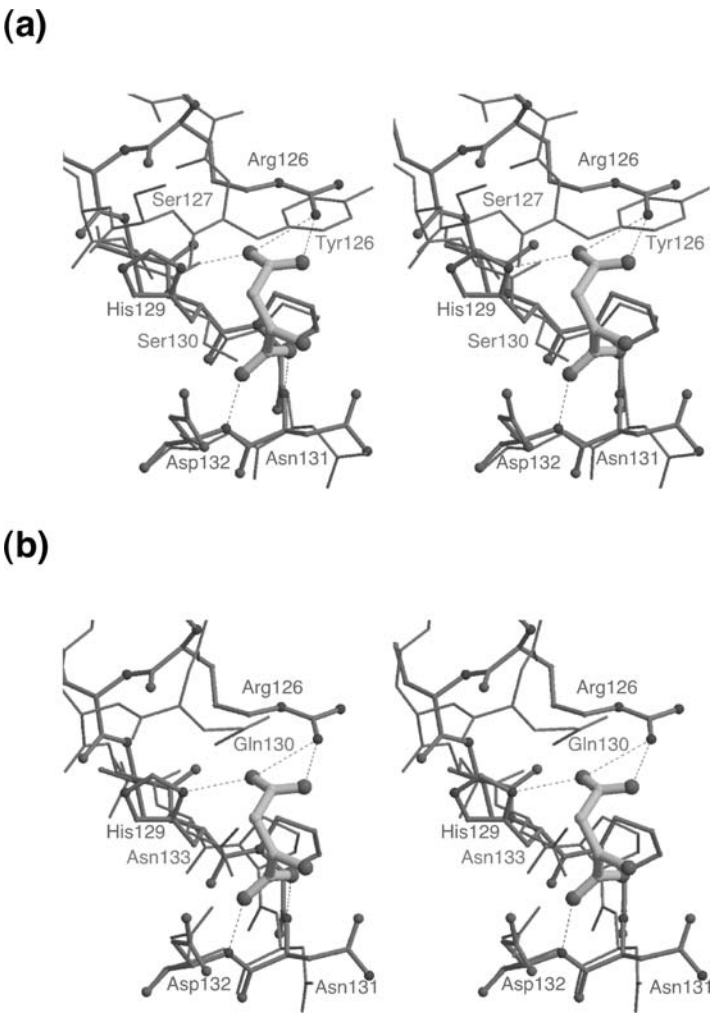


Figure 7. Superimposed comparisons of the active site and activator binding site regions of *E. coli* aspartase, *E. coli* fumarase and *Bacillus* sp. YM55-1 aspartase reprinted from Fujii *et al.* (2003). In both panels the docked L-aspartate molecule is shown by a thick ball-and-stick model. (a) Superposition of *Bacillus* sp. YM55-1 aspartase (wire model) with *E. coli* fumarase (thin ball-and-stick model). Note the extension of Tyr126 into the activation site region. (b) Superposition of *E. coli* aspartase (wire model) with *E. coli* fumarase (thin ball-and-stick model)

4. MOLECULAR ENGINEERING OF ASPARTASE AND INDUSTRIAL APPLICATIONS

In industry, aspartase is used mainly in the enzyme-mediated production of L-aspartate, an important starting compound for generating food additives and

artificial sweeteners. A survey of patents awarded for the application of aspartase reveals a number of interesting methods that have been used to alter the activity of the enzyme with a view toward its utilization in various industrial processes. In this section, we summarize the numerous methods that have been used in the industrial application of aspartase, and also provide some hints toward future efforts.

4.1. C-terminal Deletion of Amino Acid Residues

Early studies on aspartase discovered that the specific activity of this enzyme could be enhanced by brief treatment with certain proteases such as trypsin or subtilisin (Yumoto *et al*, 1980; Yumoto *et al*, 1982). Analysis showed that C-terminal truncation was responsible for this enhancement of activity. Jayasekera and co-workers (Jayasekera *et al*, 1997) undertook a detailed analysis of C-terminal deletion mutants of *E. coli* aspartase and found that a mutant aspartase truncated after Tyr 472 displayed a greater than two-fold enhancement of k_{cat} and a minimal increase in the K_M, yielding an enzyme with enhanced activities. The results of this study have been patented for use in industrial applications (US Patent 5,993,807). A similar patent (US Patent 6,015,704) was later awarded to Tsai and co-workers who combined the effects of C-terminal truncation with the substitution of selected histidine residues by glutamine (His 25, 123, 421 or 463). These two patents open the way toward the creation of novel aspartase proteins with improved enzymatic activities suitable for industrial applications.

4.2. Random Mutagenesis and Directed Evolution

The difficulties in identifying the specific residues that are important for aspartase function have led to the utilization of random mutagenesis as a viable method for obtaining aspartases having enhanced activities. Zhang and co-workers successfully used this method to obtain a mutant *E. coli* aspartase whose specific activity was increased 4-fold and which also showed slightly improved thermostability (Zhang *et al*, 1993). The characteristics of these enzymes were compared to those of another mutant aspartase produced by site-directed mutagenesis, Lys126Arg. Although both enzymes displayed similar specific activities, it is not known whether the enzyme derived from random mutagenesis is similar or identical to the Lys126Arg mutant. Another enhanced aspartase enzyme was obtained using a directed evolution approach comprising random mutagenesis followed by gene shuffling (Wang *et al*, 2000). The enzyme obtained had the following amino acid substitutions; Asn217Lys, Thr233Arg and Val367Gly. Alteration of these three amino acids resulted in an enzyme with a k_{cat}/K_M that was increased 28-fold compared to wild type, a K_M that was decreased 4.6 fold, and improved thermostability such that 61% of its initial activity was retained after a 30 min incubation at 50 °C. Under similar conditions the wild type enzyme only retained 17% of its activity. This improved aspartase would be an ideal candidate for application in industry.

4.3. Isolation of Thermostable Aspartases from Thermophilies and Psychrophiles

As outlined above, the aspartase isolated from *Bacillus* sp. YM55-1 possesses many characteristics that are desirable for industrial applications, including high specific activity and thermostability. Other examples of thermostable aspartases have also been reported. Kimura and co-workers have been awarded a patent (US Patent 4,391,910) outlining their discovery of thermophilic aspartases isolated from various *Bacillus* species such as *Bacillus aminogenes* and *Bacillus thermoaminophilus*. The enzymes described in the patent are characterized by very high specific activity, much like that of the aspartase from *Bacillus* sp. YM55-1. However, unlike the latter, the aspartases described in the patent are activated by the presence of divalent cations such as magnesium ion. This interesting difference in enzymatic characteristics may suggest that the finer characteristics of aspartases may differ considerably even among members of the same genus.

Recently, the isolation of a novel thermostable aspartase from a marine psychrophile, *Cytophaga* sp. KUC-1 (Kazuoka *et al*, 2003) has been reported whose optimal temperature of growth is 15 °C. This aspartase displays enzymatic characteristics that are similar to those of the *E. coli* enzyme, such as activation by Mg^{2+} and L-aspartate. However, the thermostability of the *Cytophaga* enzyme is considerably higher: about 80% of its activity is retained after a 60 min incubation at 50 °C. Thus the enzyme from *Cytophaga* KUC-1 represents an interesting example of an aspartase whose characteristics are a composite of those derived from aspartases from mesophilic and thermophilic sources. A comparative study of these three enzyme types should be very enlightening as regards the elucidation of the various molecular mechanisms which underlie the functional aspects of aspartase.

4.4. Immobilization of Aspartase and Aspartase-producing Bacteria within Solid Gel Matrices

Industrial production of L-aspartate using aspartase and aspartase-producing organisms is an excellent example of the successful application of an enzymatic activity in industrial processes. Since the early 1970's various methods have been developed to utilize the activity of aspartase to produce L-aspartate in bulk quantities (Sato and Tosa, 1993). Early attempts involved the immobilization of purified aspartase enzyme in polyacrylamide lattices (Tosa *et al*, 1973) and was followed by immobilization of aspartase-producing bacteria themselves in gel matrices such as polyurethane (Fusee *et al*, 1981) and polyacrylamide. The immobilization of living cells in gel matrices provides certain advantages over immobilizing pure enzymes such as enhanced stability of the enzymatic activity and the saving of labour costs associated with enzyme purification. Future applications in this direction may involve utilizing bacteria that produce enhanced forms of the aspartase, for example *E. coli* cells expressing the aspartase from *Bacillus* sp. YM55-1.

4.5. Other Studies

A number of interesting studies on aspartase provide hints toward future work on applications for this enzyme. Early studies by Murase and co-workers have shown that aspartase is partially active in a dimeric state induced by the addition of a denaturant (Murase *et al*., 1993). In support of this finding, recent work by Kong and co-workers succeeded in constructing a novel form of aspartase in which two original subunits of *E. coli* aspartase were joined by a suitable peptide linker to produce a dimer-mimicking, monomeric form of the enzyme that was active (Kong *et al*., 2002). This monomeric form of aspartase displayed improved stability at high temperatures compared to the wild type enzyme. Such approaches toward stabilizing the structure and ultimately the activity of aspartase represent a novel alternate method for improving the high efficiency of this enzyme.

Another interesting approach that has been tested involves the construction of hybrid and/or chimeric proteins. By ingenious utilization of gene fusion and *in vivo* selection for the desired activity Sheng *et al*. succeeded in constructing a 74-kDa polypeptide hybrid enzyme combining the activities of α-aspartyl dipeptidase (85% of wild type) and aspartase (87% of wild type) into a single polypeptide (Sheng *et al*., 2005). As an added bonus, this novel hybrid enzyme displayed improved thermostabilities with regard to both activities. This astonishing example of protein engineering is indicative of the potential that still remains to be tapped in aspartase. We too have recently begun to study various chimeric constructs of aspartase genes with a view to generating novel aspartases having improved thermostabilities and enzymic activities.

5. CONCLUDING REMARKS

The accumulation of knowledge on the structure and function of aspartase provides us with a detailed view of the mechanisms that underlie the unique functional aspects of this intriguing enzyme. Nevertheless, experimentation on aspartase never fails to produce exciting new developments that promise to lead to previously unsuspected applications and novel concepts. Exploration of the functional and structural aspects of aspartase is work in progress, and future studies will further deepen our understanding of this enzyme.

REFERENCES

Acuna, G., Ebeling, S. and Hennecke, H. (1991). Cloning, sequencing, and mutational analysis of the *Bradyrhizobium japonicum* fumC-like gene: evidence for the existence of two different fumarases. J Gen Microbiol *137*, 991–1000.

Beeckmans, S. and Van Driessche, E. (1998). Pig heart fumarase contains two distinct substrate binding sites differing in affinity. J Biol Chem *273*, 31661–31669.

Blanchard, J.S. and Cleland, W.W. (1980). Use of isotope effects to deduce the chemical mechanism of fumarase. Biochemistry *19*, 4506–4513.

Burland, V., Plunkett, G., 3rd, Sofia, H.J., Daniels, D.L. and Blattner, F.R. (1995). Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92.8 through 100 minutes. Nucleic Acids Res *23*, 2105–2119.

Chen, H.H., Chen, J.T. and Tsai, H. (1996). Site-directed mutagenesis of cysteinyl residues in aspartase of *Escherichia coli*. Ann N Y Acad Sci *799*, 70–73.

Cook, R.P. and Woolf, B. (1928). The deamination and synthesis of L-aspartic acid in the presence of bacteria. Biochem J *22*, 474–481.

Ellfolk, N. (1953a). Studies on aspartase. I. Quantitative separation of aspartase from bacterial cells, and its partial purification. Acta Chem Scand *7*, 824–830.

Ellfolk, N. (1953b). Studies on aspartase. II. On the chemical nature of aspartase. Acta Chem Scand *7*, 1155–1163.

Ellfolk, N. (1954). Studies on aspartase. III. On the specificity of aspartase. Acta Chem Scand *8*, 151–156.

Falzone, C.J., Karsten, W.E., Conley, J.D. and Viola, R.E. (1988). L-aspartase from *Escherichia coli*: substrate specificity and role of divalent metal ions. Biochemistry *27*, 9089–9093.

Fujii, T., Sakai, H., Kawata, Y. and Hata, Y. (2003). Crystal structure of thermostable aspartase from *Bacillus* sp. YM55-1: structure-based exploration of functional sites in the aspartase family. J Mol Biol *328*, 635–654.

Fusee, M.C., Swann, W.E. and Calton, G.J. (1981). Immobilization of *Escherichia coli* Cells Containing Aspartase Activity with Polyurethane and Its Application for l-Aspartic Acid Production. Appl Environ Microbiol *42*, 672–676.

Guest, J.R., Roberts, R.E. and Wilde, R.J. (1984). Cloning of the aspartase gene (aspA) of *Escherichia coli*. J Gen Microbiol *130*, 1271–1278.

Ida, N. and Tokushige, M. (1984). Chemical modification of essential histidine residues in aspartase with diethylpyrocarbonate. J Biochem (Tokyo) *96*, 1315–1321.

Ida, N. and Tokushige, M. (1985). Assignment of catalytically essential cysteine residues in aspartase by selective chemical modification with N-(7-dimethylamino-4-methylcoumarynyl) maleimide. J Biochem (Tokyo) *98*, 793–797.

Ida, N. and Tokushige, M. (1985). L-Aspartate-induced activation of aspartase. J Biochem (Tokyo) *98*, 35–39.

Jayasekera, M.M., Saribas, A.S. and Viola, R.E. (1997). Enhancement of catalytic activity by gene truncation: activation of L-aspartase from *Escherichia coli*. Biochem Biophys Res Commun *238*, 411–414.

Jayasekera, M.M., Shi, W., Farber, G.K. and Viola, R.E. (1997). Evaluation of functionally important amino acids in L-aspartate ammonia-lyase from *Escherichia coli*. Biochemistry *36*, 9145–9150.

Jayasekera, M.M. and Viola, R.E. (1999). Recovery of catalytic activity from an inactive aggregated mutant of L-aspartase. Biochem Biophys Res Commun *264*, 596–600.

Karsten, W.E., Gates, R.B. and Viola, R.E. (1986). Kinetic studies of L-aspartase from *Escherichia coli*: substrate activation. Biochemistry *25*, 1299–1303.

Karsten, W.E., Hunsley, J.R. and Viola, R.E. (1985). Purification of aspartase and aspartokinase-homoserine dehydrogenase I from *Escherichia coli* by dye-ligand chromatography. Anal Biochem *147*, 336–341.

Karsten, W.E. and Viola, R.E. (1991). Kinetic studies of L-aspartase from *Escherichia coli*: pH-dependent activity changes. Arch Biochem Biophys *287*, 60–67.

Katoh, K., Kuma, K., Toh, H. and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res *33*, 511–518.

Kawata, Y., Tamura, K., Kawamura, M., Ikei, K., Mizobata, T., Nagai, J., Fujita, M., Yano, S., Tokushige, M. and Yumoto, N. (2000). Cloning and over-expression of thermostable *Bacillus* sp. YM55-1 aspartase and site-directed mutagenesis for probing a catalytic residue. Eur J Biochem *267*, 1847–1857.

Kawata, Y., Tamura, K., Yano, S., Mizobata, T., Nagai, J., Esaki, N., Soda, K., Tokushige, M. and Yumoto, N. (1999). Purification and characterization of thermostable aspartase from *Bacillus* sp. YM55-1. Arch Biochem Biophys *366*, 40–46.

Kazuoka, T., Masuda, Y., Oikawa, T. and Soda, K. (2003). Thermostable aspartase from a marine psychrophile, *Cytophaga* sp. KUC-1: molecular characterization and primary structure. J Biochem (Tokyo) *133*, 51–58.

Kong, X., Li, Z., Gou, X., Zhu, S., Zhang, H., Wang, X. and Zhang, J. (2002). A monomeric L-aspartase obtained by *in vitro* selection. J Biol Chem *277*, 24289–24293.

Lee, T.T., Worby, C., Bao, Z.Q., Dixon, J.E. and Colman, R.F. (1999). His68 and His141 are critical contributors to the intersubunit catalytic site of adenylosuccinate lyase of Bacillus subtilis. Biochemistry *38*, 22–32.

Mizuta, K. and Tokushige, M. (1975). Studies on aspartase. II. Role of sulfhydryl groups in aspartase from *Escherichia coli*. Biochim Biophys Acta *403*, 221–231.

Murase, S., Kawata, Y. and Yumoto, N. (1993). Identification of an active dimeric form of aspartase as a denaturation intermediate. J Biochem (Tokyo) *114*, 393–397.

Nuiry, II, Hermes, J.D., Weiss, P.M., Chen, C.Y. and Cook, P.F. (1984). Kinetic mechanism and location of rate-determining steps for aspartase from *Hafnia alvei*. Biochemistry *23*, 5168–5175.

Porter, D.J. and Bright, H.J. (1980). 3-Carbanionic substrate analogues bind very tightly to fumarase and aspartase. J Biol Chem *255*, 4772–4780.

Quastel, J.H. and Woolf, B. (1926). The equilibrium between L-aspartic acid, fumaric acid and ammonia in the presence of resting bacteria. Biochem J 20, 545–555.

Rose, I.A. and Weaver, T.M. (2004). The role of the allosteric B site in the fumarase reaction. Proc Natl Acad Sci U S A *101*, 3393–3397.

Rudolph, F.B. and Fromm, H.J. (1971). The purification and properties of aspartase from *Escherichia coli*. Arch Biochem Biophys *147*, 92–98.

Saribas, A.S., Schindler, J.F. and Viola, R.E. (1994). Mutagenic investigation of conserved functional amino acids in *Escherichia coli* L-aspartase. J Biol Chem *269*, 6313–6319.

Sato, T. and Tosa, T. (1993). Production of L-aspartic acid. Bioprocess Technol *16*, 15–24.

Schindler, J.F. and Viola, R.E. (1994). Mechanism-based inactivation of L-aspartase from *Escherichia coli*. Biochemistry *33*, 9365–9370.

Sheng, Y., Li, S., Gou, X., Kong, X., Wang, X., Sun, Y. and Zhang, J. (2005). The hybrid enzymes from alpha-aspartyl dipeptidase and L-aspartase. Biochem Biophys Res Commun *331*, 107–112.

Shi, W., Dunbar, J., Jayasekera, M.M., Viola, R.E. and Farber, G.K. (1997). The structure of L-aspartate ammonia-lyase from *Escherichia coli*. Biochemistry *36*, 9136–9144.

Sun, D.X. and Setlow, P. (1991). Cloning, nucleotide sequence, and expression of the *Bacillus subtilis* ans operon, which codes for L-asparaginase and L-aspartase. J Bacteriol *173*, 3831–3845.

Takagi, J.S., Ida, N., Tokushige, M., Sakamoto, H. and Shimura, Y. (1985). Cloning and nucleotide sequence of the aspartase gene of *Escherichia coli* W. Nucleic Acids Res *13*, 2063–2074.

Takagi, J.S., Tokushige, M. and Shimura, Y. (1986). Cloning and nucleotide sequence of the aspartase gene of *Pseudomonas fluorescens*. J Biochem (Tokyo) 100, 697–705.

Tosa, T., Sato, T., Mori, T., Matuo, Y. and Chibata, I. (1973). Continuous production of L-aspartic acid by immobilized aspartase. Biotechnol Bioeng *15*, 69–84.

Viola, R.E. (2000). L-aspartase: new tricks from an old enzyme. Adv Enzymol Relat Areas Mol Biol *74*, 295–341.

Virtanen, A.I. and Tarnanen, J. (1932). Die Enzymatische Spaltung und Synthese der Asparaginsaure. Biochem Z *250*, 193–211.

Wang, L.J., Kong, X.D., Zhang, H.Y., Wang, X.P. and Zhang, J. (2000). Enhancement of the activity of L-aspartase from *Escherichia coli* W by directed evolution. Biochem Biophys Res Commun *276*, 346–349.

Weaver, T. and Banaszak, L. (1996). Crystallographic studies of the catalytic and a second site in fumarase C from *Escherichia coli*. Biochemistry *35*, 13955–13965.

Weaver, T., Lees, M. and Banaszak, L. (1997). Mutations of fumarase that distinguish between the active site and a nearby dicarboxylic acid binding site. Protein Sci *6*, 834–842.

Weaver, T.M., Levitt, D.G., Donnelly, M.I., Stevens, P.P. and Banaszak, L.J. (1995). The multisubunit active site of fumarase C from *Escherichia coli*. Nat Struct Biol *2*, 654–662.

Woods, S.A., Miles, J.S. and Guest, J.R. (1988). Sequence homologies between argininosuccinase, aspartase, and fumarase: a family of structurally-related enzymes. FEMS Microbiol Lett *51*, 181–186.

Wu, C.Y., Lee, H.J., Wu, S.H., Chen, S.T., Chiou, S.H. and Chang, G.G. (1998). Chemical mechanism of the endogenous argininosuccinate lyase activity of duck lens delta2-crystallin. Biochem J *333*(Pt 2), 327–334.

Yoon, M.Y., Thayer-Cook, K.A., Berdis, A.J., Karsten, W.E., Schnackerz, K.D. and Cook, P.F. (1995). Acid/base chemical mechanism of aspartase from *Hafnia alvei*. Arch Biochem Biophys *320*, 115–122.

Yumoto, N., Mizuta, K., Tokushige, M. and Hayashi, R. (1982). Studies on aspartase VIII. Protease-mediated activation: comparative survey of protease specificity for activation and peptide cleavage. Physiol Chem Phys *14*, 391–397.

Yumoto, N., Okada, M. and Tokushige, M. (1982). Biospecific inactivation of aspartase by L-aspartic-beta-semialdehyde. Biochem Biophys Res Commun *104*, 859–866.

Yumoto, N., Tokushige, M. and Hayashi, R. (1980). Studies on aspartase. VI. Trypsin-mediated activation releasing carboxy-terminal peptides. Biochim Biophys Acta *616*, 319–328.

Zhang, H.Y., Zhang, J., Lin, L., Du, W.Y. and Lu, J. (1993). Enhancement of the stability and activity of aspartase by random and site-directed mutagenesis. Biochem Biophys Res Commun *192*, 15–21.

CHAPTER 32

TRANSGLUTAMINASES

MARÍA JESÚS ARRIZUBIETA*

Departamento de Ciencias del Medio Natural, Universidad Pública de Navarra, Pamplona, Spain
**mjesus.arrizubieta@unavarra.es*

1. INTRODUCTION

Transglutaminases (TG: EC 2.3.2.13, protein glutamine:amine γ-glutamyl-transferase) are a group of thiol enzymes that catalyse the post-translational modification of proteins mainly by protein to protein cross-linking, but also through the covalent conjugation of polyamines, lipid esterification, or the deamidation of glutamine residues (Folk and Chung, 1973; Lorand and Conrad, 1984; Aeschlimann and Paulsson, 1994; Nemes *et al*, 1999). Transglutaminases are widely distributed among bacteria, plants and animals. Mammalian transglutaminases comprise a group of structurally and phylogenetically related multidomain enzymes strictly dependent on calcium for their activity (Grenard *et al*, 2001). These enzymes are related to different physiological processes and diseases and consequently have raised medical and pharmacological interests in their potential as therapeutic targets or diagnostic aids (Griffin *et al*, 2002). In contrast to mammalian transglutaminases, the microbial transglutaminase from *Streptomyces* is a monomeric enzyme that does not require calcium for activity. Phylogenetically this enzyme is not related to mammalian transglutaminases and presents a novel catalytic mechanism (Kashiwagi *et al*, 2002). Its discovery has enabled a diversity of industrial applications in the food and textile industries. Transglutaminase activity has also been found in higher and lower plants (Del Duca and Serafini-Fracassini, 2005), and the available data indicate that plant transglutaminases are similar in overall structure and catalytic mechanism to those of mammals (Villalobos *et al*, 2004).

2. ENZYME ACTIVITY AND MECHANISM OF CATALYSIS

TGs catalyse the acyl-transfer reaction between the γ-carboxyamide group of a peptide-bound glutamine residue and a primary amine. In this reaction the glutamine side chain serves as the acyl donor, whereas the primary amine functions as acceptor.

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 567–581

The outcome is the generation of a covalent bond between the two substrates and the release of ammonia. In most cases described, the primary amine is the ε-amino group of a lysine residue, and the reaction results in the formation of ε-(γ-glutamyl)lysine linkages. Depending on the nature of the substrates (proteins or peptides), the reaction may result in protein cross-linking or peptide conjugation to proteins. When primary amines are not available, water can function as the acyl group acceptor, with the consequent deamidation of the glutamine residue (Folk and Chung, 1973; Lorand and Conrad, 1984). In addition, the hydroxyl groups of hydroxyceramides can also function as acyl acceptor substrates resulting in ester bond formation between the glutamine side chains of proteins and hydroxylipids (Nemes *et al*., 1999).

The transglutaminase reaction is reversible and proceeds via a modified double-displacement mechanism in which an acyl-enzyme intermediate is formed between the acyl portion of the glutamine substrate and the sulfhydryl group of the catalytically active cysteine residue on the enzyme. An acyl acceptor, like water or a primary amine, reacts with this intermediate resulting in deamidation or cross-linking reactions, respectively (Folk, 1969; Folk and Chung, 1973). The first crystal structure of a TG, that of blood coagulation factor XIII (FXIII), revealed striking similarities to cysteine-proteases at the active site, leading to the proposal of a similar catalytic mechanism for the TG-catalysed cross-linking reaction (Pedersen *et al*., 1994). According to this mechanism, Cys-314 acts as the nucleophile and His-373 as the acid/base catalyst. Asp-396 is proposed to play a secondary role, stabilizing the protonated form of His-373 and providing a favourable orientation of this residue. Based on the structural and mechanistic similarities between FXIII-like TGs and cysteine-proteases, it has been proposed that these two groups of enzymes, catalysing almost opposite reactions, evolved from a common ancestor (Pedersen *et al*., 1994).

TGs are synthesized as inactive zymogens that require activation before exhibiting their transamidating activity (Lorand and Conrad, 1984). The mechanisms for controlling the transamidating activity varies from enzyme to enzyme and involves activation by proteolytic cleavage, calcium binding, GTP binding, substrate sequestration, and often combinations of two or more of these mechanisms (Lorand and Conrad, 1984; Griffin *et al*., 2002).

Different enzymes display specificity differences towards substrate proteins. This constitutes an important aspect of their biological function and it is also relevant to their biotechnological applications. In general, there is more stringent specificity for the glutamine substrate whereas the specificity for the primary amine substrate is broader. The nature of the amino acids surrounding the sensitive glutamine residues and the degree of the latters' exposure to solvent appear to be the main determinants of substrate specificity (Coussons *et al*., 1992).

3. ENZYME TYPES

3.1. Mammalian Transglutaminases

Nine different transglutaminase encoding genes have been characterized in mammals (Ichinose *et al*., 1990; Aeschlimann and Paulsson, 1994; Aeschlimann

et al., 1998). The products of eight of them are active enzymes whereas the remaining one is a non-catalytic protein, the erythrocyte protein band 4.2. These genes present a high degree of sequence similarity and share the same genetic organization.

All characterized mammalian TGs bind calcium and function as metal-enzyme complexes in different biological processes (Aeschlimann and Paulsson, 1994). Blood coagulation factor XIII catalyses covalent cross-linking between fibrin molecules during blood clot formation (Pisano *et al*., 1968; Ichinose *et al*., 1990). There are six different TGs expressed during epidermal differentiation. These enzymes cross-link the structural proteins that constitute the protective cornified cell envelop of differentiating keratinocytes (Nemes and Steinert, 1999; Candi *et al*., 2005). Tissue transglutaminase, or TG2, besides being a cross-linking enzyme functions also as a GTPase, and participates in processes as diverse as apoptosis, development, intracellular signalling, cell-matrix interactions and cell migration (Akimov *et al*., 2000; Piacentini *et al*., 2000; Mangala *et al*., 2005; Sarang *et al*., 2005). Prostate transglutaminase, or TG4, is an extracellular TG secreted by the prostate gland into the seminal fluid and is responsible for the production of the rodent vaginal plug (Aeschlimann and Paulsson, 1994).

3.2. Microbial Transglutaminases

The first characterized microbial transglutaminase (MTG) was that of the bacterium *Streptomyces mobaraensis* (previously termed *Streptoverticillium mobaraensis*) (Ando *et al*., 1989). This enzyme is secreted as a zymogen that is sequentially processed by two endogenous enzymes to yield the mature form (Zotzel *et al*., 2003). The mature enzyme is a monomeric protein 331 amino acids long and contains a single cysteine (Cys-64) which is the catalytic residue (Kanaji *et al*., 1993; Pasternack *et al*., 1998). The amino acid sequence of MTG bears little significant similarity to the FXIII-like TGs or to any other known sequence. MTG does not require calcium for activity, shows broad substrate specificity and can be produced at relatively low cost. These properties are advantageous for industrial applications. Currently the enzyme is produced by fermentation using a *Streptomyces* strain; however, a lot of effort has been directed towards achieving production of recombinant MTG in a heterologous host (Yokoyama *et al*., 2004).

TG activity has also been found in *Bacillus* species (Kobayashi-K *et al*., 1996). The enzyme purified from *B. subtilis* catalyses the cross-linking of the spore coat protein GerQ thereby contributing to the physical and chemical resistance of the bacterial spore (Ragkousi and Setlow, 2004). In addition, bacterial virulence factors such as *Escherichia coli* cytotoxic necrotizing factor and *Bordetella* dermonecrotic toxin present transglutaminase activity. They exert their cytotoxic effects by activating the small intracellular GTPases of the Rho family by

deamidation of specific Gln residues or by cross-linking the GTPases with polyamines (Horiguchi, 2001).

3.3. Plant Transglutaminases

Plant transglutaminases are present in various compartments such as chloroplasts, mitochondria, cell walls and cytoplasm. Their functions are related to plant growth, cell division, differentiation, programmed cell death, fertilization and stress (Serafini-Fracassini *et al*., 1995). So far only three TG have been characterized at the molecular level, one in *Arabidopsis* (Della Mea *et al*., 2004) and two proteins of 39 and 58 kDa in the chloroplast of *Zea mays* (Della Mea *et al*., 2004; Villalobos *et al*., 2004). The expression and activity of the corn enzymes is light dependent and it has been proposed that they might play a role in the light-harvesting process in thylakoids (Della Mea *et al*., 2004).

4. ENZYME STRUCTURE

4.1. Factor XIII

The structure of the zymogen of the A subunit of factor XIII (FXIIIA) has been published (Yee *et al*., 1994). The solved structures correspond to a homodimer. The region corresponding to the mature enzyme is composed of four domains. Starting from the N-terminus, there is a β-sandwich domain (residues 43–184) followed by a catalytic core (residues 185–515) and two β-barrels - barrel 1 (residues 516–628)

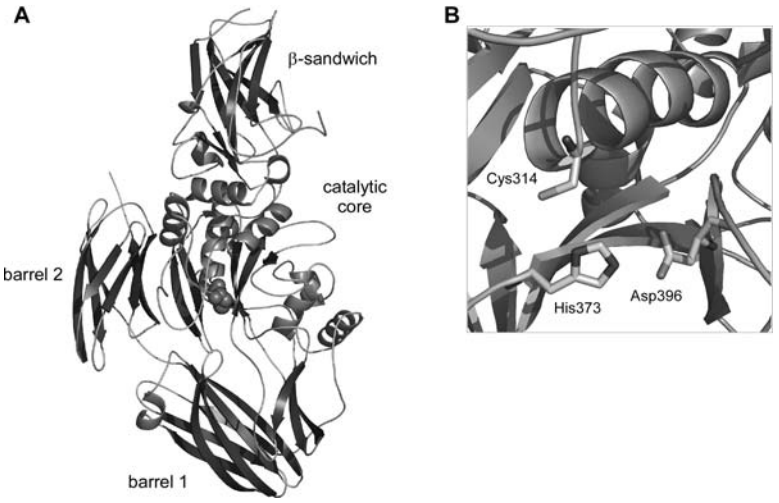


Figure 1. A) 3D structure of factor XIIIA. Only one subunit of the dimer is represented. The catalytic cysteine is represented in space-filling mode. B) Detail of the catalytic triad

and barrel 2 (residues 629–727) - at the C-terminus (Fig. 1). The core domain includes 10 segments of α-helix arranged around a twisted β-sheet of six antiparallel strands. The two C-terminal barrels have a seven-stranded β-barrel fold, with a fibronectin-like structure. Residues Cys-314, His-373 and Asp-396 constitute the catalytic triad of factor XIIIA and they are located at the base of a cavity bounded by the core and barrel 1 domains. The catalytically active cysteine, Cys-314, sits at the N-terminus of a long helix in the core domain and is hydrogen bonded to His-373, which in turn forms a hydrogen bond with Asp-396. Both His-373 and Asp-396 sit on two β-strands of the same sheet. In addition, Cys-314 is also hydrogen bonded to Tyr-560 located in one of the loops of the barrel 1 domain.

In the zymogen structure the active site of the molecule is inaccessible to solvent, and thus to substrate, due to both intra- and inter-subunit interactions. Access to the catalytic Cys-314 residue is blocked by Tyr-560, and movement of the β-barrel 1 domain is required to allow the substrate to approach the nucleophilic Cys-314. In addition, the structure of the dimer shows that the activation peptide of each A-subunit crosses the dimer interface and partially occludes the opening of the active site in the catalytic core of the other subunit. Only after removal of this pro-peptide is the active site accessible to substrates. Crystallographic studies have shown that neither calcium binding (Fox *et al*., 1999) nor pro-peptide cleavage alone (Yee *et al*., 1995) causes the large conformational changes required to expose the active site. Based on these results it has been postulated that binding of both Ca^{2+} and the acyl donor substrate is required to trigger the conformational change necessary to expose the active site of the enzyme.

4.2. Other Animal Transglutaminases

Besides human FXIIIA, the X-ray structures of three other animal TGs have been solved: human tissue transglutaminase, TG2 (Liu *et al*., 2002), human epithelial transglutaminase, TG3 (Ahvazi *et al*., 2002) and the fish-derived TG, fTG (Noguchi *et al*., 2001). All these TGs present the same overall four domain structure as FXIIIA: β-sandwich, α/β catalytic core, barrel 1 and barrel 2. The position of the catalytic triad Cys-His-Asp is conserved in the three enzymes, and in all of them the active site is shielded from contact with the solvent thus requiring major structural changes for enzyme activation. The conservation of these features indicates a common mechanism of catalysis and a common phylogenetic origin for the members of this family of enzymes. Nevertheless, the solved structures of these additional TGs have revealed the structural basis for the biochemical peculiarities of these enzymes, such as the GTP regulation of TG2 activity (Liu *et al*., 2002) and the mechanism of calcium activation of TG3 (Ahvazi *et al*., 2002).

4.3. *Streptomyces mobaraensis* Transglutaminase

The crystal structure of the microbial transglutaminase from *Streptomyces mobaraensis* (Kashiwagi *et al*., 2002) revealed a completely different structure

from that of FXIIIA and related TGs. Furthermore, the structure of MTG seems to present a novel fold since it does not resemble any structure published to date. The configuration of the catalytic residues also suggests a distinct catalytic mechanism that might account for the peculiar biochemical characteristics of this enzyme.

In contrast to the multidomain structure of the FXIII-like TGs, MTG presents a single domain having a disk-like shape (Fig. 2). The active site is located at the bottom of a deep cleft at the edge of the disk. The structure of the MTG belongs to the $\alpha+\beta$ folding class, containing 11 α-helices and 8 β-strands. These secondary structure elements are arranged so that a central β-sheet with seven antiparallel strands is surrounded by the α-helices.

The size of MTG (331 amino acids) is similar to that of the core domain of FXIIIA (330 amino acids). Although they both belong to the $\alpha+\beta$ folding class, their overall folding patterns are considerably different. Nevertheless, the arrangements of the secondary structure elements around their active sites are very similar. Both active site cysteines occur near the N-termini of α-helices, and in each TG this α-helix is flanked by a four-stranded β-sheet that contains the other two residues of the catalytic triad.

Unlike the catalytically active cysteine groups of animal TGs, the sulfhydryl group of Cys-64 is partially exposed to the solvent and can readily react with substrates, which may explain the higher reaction rate of MTG. Besides Cys-64, residues Asp-255 and His-274 constitute the catalytic triad. The most striking difference between MTG and the FXIII-like TGs is that the positions occupied by Asp-255 and His-274 in MTG superimpose well on the positions of His-373 and Asp-396 in FXIIIA, respectively. In other words, in the MTG molecule the relative positions of the catalytically important His and Asp residues are reversed

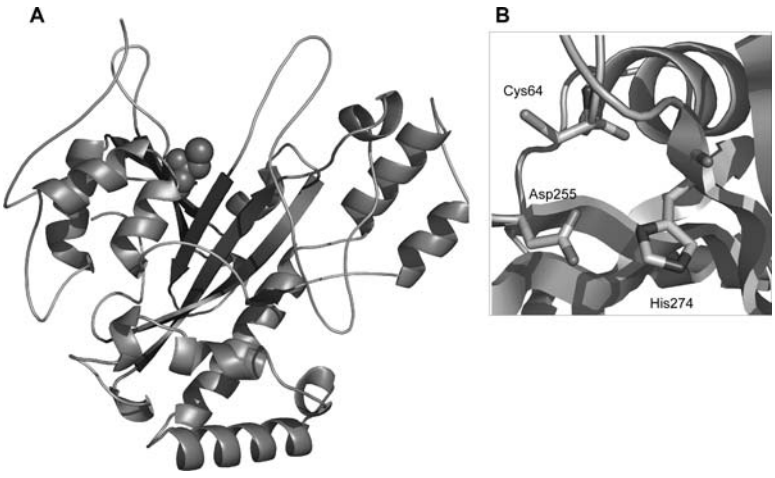


Figure 2. A) 3D structure of the microbial TG from *Streptomyces mobaraensis*. The catalytic cysteine is represented in space-filling mode. B) Detail of the catalytic triad

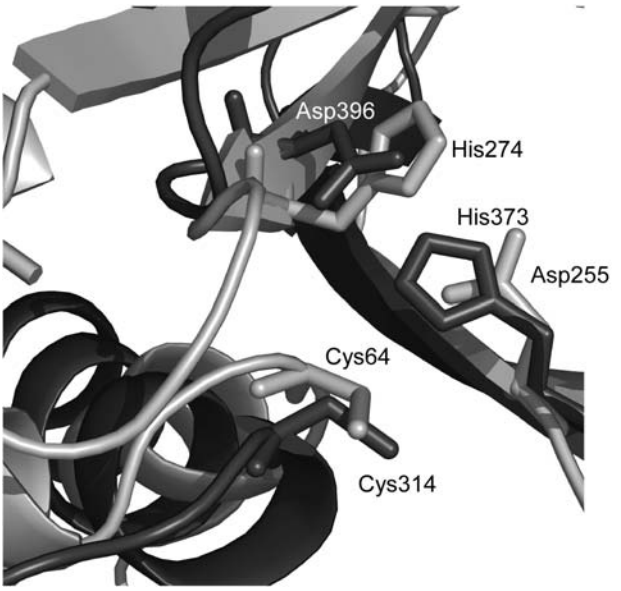


Figure 3. Comparison of the active centres of factor XIIIA (dark grey) and MTG (light grey). Organization of secondary structure elements and catalytic residues at the active site

relative to the Cys residue (Fig. 3). These results have been taken to indicate that in the catalytic mechanism of MTG, Asp-255, and not His-274, acts as the acid/base catalyst. It has been suggested that the negatively charged state of Asp-255 during the reaction might be responsible for the weaker deamidation activity of MTG compared to the FXIII-like TGs since this would favour positively charged groups acting as acyl acceptors instead of neutral species such as water.

The similarities between the active site structures and the differences in the overall structures between MTG and the FXIII-like TGs suggest that the relationship between these enzymes is a special case of convergent molecular evolution.

5. INDUSTRIAL APPLICATIONS OF TRANSGLUTAMINASES

TG-mediated cross-linking of proteins has dramatic effects on their physical and chemical properties. This has triggered the use of these biocatalysts in a wide range of industrial sectors, from cosmetics to the food industry. In addition, and arising from the involvement of this group of enzymes in many physiological and pathological processes, they have found important applications in the pharmaceutical industry. Their biotechnological potential is best reflected in the rapidly growing number of patent applications regarding this group of enzymes (Griffin *et al*., 2002).

The first assays to test possible industrial uses of TG were carried out to modify food proteins using mammalian enzymes (Matheis and Whitaker, 1987). Despite the interest in these enzymes for modulating food rheological properties, limited

supply hindered their commercial utilization. The situation changed in 1989 with the purification of the microbial TG from *Streptomyces mobaraensis* (Ando *et al*, 1989) which could be produced relatively cheaply by fermentation methods. Subsequently, the application of TG in industry, and especially in the food industry, began to increase at a steady rate. It is remarkable that sixteen years later, the microbial transglutaminase obtained from *Streptomyces* is still the only commercial source of MTG. The use of microbial transglutaminase in food processing has been covered in a series of reviews (Nielsen, 1995; Motoki and Seguro, 1998; Kuraishi *et al*, 2001; Yokoyama *et al*, 2004).

5.1. Meat and Fish Products

Muscle proteins, especially myosin, are polymerised by TG (Huang *et al*, 1992) and the resulting cross-links strengthen the protein network in muscle-derived products. The addition of the TG during the setting process of surimi results in a gel with greater breaking stress and improved elasticity (Sakamoto *et al*, 1995), both of which are principle determinants of the value of the product. Fish species with high proteolytic activity were traditionally not capable of being used for surimi production. In these under utilized fish species the addition of TG has been reported to improve the gel characteristics of fish paste (Haejung *et al*, 1996) thereby allowing their utilization in surimi production.

In the case of sausages, hams and other meat products the effect of TG addition is improved texture with higher breaking strength or firmness, and increased deformation and elasticity of the gel. This leads to an increase in manufacturers' profits due to reduced product loss during processing and slicing. On the other hand, the use of TG results in the production of reduced salt/phosphate meat products with improved water-holding capacity and texture, which constitutes a valuable health benefit for the consumer. The covalent nature of the isopeptide bonds created by TG determines that the changes in structure are resistant to both heating at high temperatures and freezing. Thus, manufacturing of canned or frozen products benefits from the utilization of this enzyme. One of the most successful applications of TG in the meat industry is probably the elaboration of restructured meat products. Although other methods had been previously described and used for this purpose, the combination of TG and caseinate has emerged as a very convenient and practical method to obtain restructured meat products of high quality (Motoki and Seguro, 1998; Kuraishi *et al*, 2001).

5.2. Dairy Industry

Of the proteins present in milk, caseins, which have an open conformation, react readily with TG, while globular whey proteins react only under conditions that favour unfolding (Matsumura *et al*, 1996; Lorenzen *et al*, 1998). In yoghurt manufacturing, TG treatment of milk results in an increase in firmness and viscosity and reduced syneresis. In cheese manufacturing, TG treatment of milk results in an

increase in curd yield, a less dry texture and reduced whey separation. Processed cheese products treated with the enzyme present higher heat stability, maintaining viscosity when melting. In ice-cream manufacturing, TG application allows the production of low calorie, sugar-free, ice-cream which is softer, smoother and easier to scoop (Motoki and Seguro, 1998; Kuraishi *et al*, 2001). TG also has potential commercial applications as a food-grade additive capable of improving the heat stability of milk (Osullivan *et al*, 2002), thus conferring resistance to coagulation at sterilization temperatures or gelation during storage. Enzymatic cross-linking has also potential applications in controlling the stability of milk protein-containing emulsions and foams (Dickinson and Yamamoto, 1996).

5.3. Bakery and Noodles

Gluten, the protein of wheat flour, can be cross-linked by TG and the action of the enzyme reinforces the protein network structure changing the viscoelastic properties of the dough (Larre *et al*, 2000). Consequently, pasta manufactured from TG-treated dough has a higher breaking strength, or firmness. In addition, treated pasta releases fewer solids into the boiling water, retaining its firmness and elasticity for a longer period after cooking (Kuraishi *et al*, 2001). TG has also found use in the baking industry as an improver of dough stability and loaf volume (Gerrard *et al*, 1998), as well as an improver of the lift of puff pastry and the volume of yeasted croissants (Gerrard *et al*, 2000). These effects are retained even after freezing offering a potential solution to the problem of dough deterioration during frozen storage. In addition, deep-fried dough products such as doughnuts absorb 25% less oil when the dough has been treated with transglutaminase, with the concomitant health benefit for the consumer (Kuraishi *et al*, 2001).

5.4. Soy Derivatives

TG cross-links soy globulins resulting in modification of the gelation and textural characteristics of soybean products. In the case of tofu, a soybean curd product obtained by coagulation of soybean proteins, TG treatment increases its water-holding properties and results in tofu with a smoother and firmer texture (Nonaka *et al*., 1994). In addition, treatment with the enzyme allows for better control of the coagulation reaction and a reduction in weight loss during retort cooking. Soy proteins are added to many processed foods such as sausages, ham and surimi to improve their textural and nutritional characteristics. Thus, TG offers the potential to modulate textural properties of those foods to which soy proteins are added (Motoki and Seguro, 1998).

5.5. Bioavailability of Cross-linked Proteins

The wide application of MTG in food production has raised concerns about the potential effects of the cross-linked proteins on consumer's health and the

bioavailability of the cross-linked residues. Glutamine-lysine isopeptide bonds are widely distributed in natural products and foodstuffs (with the exception of milk) and cooking itself results in glutamine-lysine isopeptide bond formation (Motoki and Seguro, 1998). Mankind has therefore been ingesting food containing glutamine-lysine isopeptide bonds since prehistoric times, suggesting that their consumption does not pose a health risk. Mammalian digestive enzymes do not cleave this isopeptide bond. However, two kinds of enzymes have been found in kidney and in intestinal brush border membranes that degrade the bond *in vivo* (Seguro *et al*, 1995; Yokoyama *et al*, 2004). Furthermore, nutritional studies have indicated that the glutamine-lysine moiety in cross-linked proteins can be metabolised and that the lysine is incorporated into proteins (Seguro *et al*, 1996). Thus, TG catalysed cross-linking does not reduce the nutritional value of the residues implicated and/or proteins. The allergenic risk of added MTG in food has been assessed and no safety concerns with regard to the allergenic potential of MTG have been identified (Pedersen *et al*, 2004).

5.6. Medical Applications

As mentioned above, TGs participate in a myriad of physiological processes. Therefore, it is not surprising that these enzymes have been implicated in different pathological states. Some of these conditions are alleviated by supplementation with exogenous TG activity but in many cases, related molecules, such as TG inhibitors or antiTG-specific antibodies, find application as therapeutic or diagnostic agents.

The first medical application of a TG was the use of FXIII in blood clotting products to control bleeding during surgery and as a general tissue adhesive. In addition, FXIII has found application in substitutive therapy of congenital or acquired deficiencies in FXIIIA that result in severe bleeding due to lack of fibrin clot stabilization. Tissue transglutaminase, or TG2, has also received commercial interest as a tissue adhesive and as a cell adhesion protein in medical implants (Collighan *et al*, 2002).

Celiac disease is a gluten sensitive enteropathy that is characterized by atrophy of the intestinal villi and results in malabsorption and malnutrition. The causative protein is gliadin, a component of wheat gluten, and the only satisfactory treatment involves a lifelong gluten-free diet. IgA auto-antibodies against endomysium are often found in sufferers and have been a useful screening method, avoiding the need for intestinal biopsy. The main endomysial auto-antigen has been shown to be the tissue TG, TG2 (Dieterich *et al*, 1997). An ELISA method has been developed based on recombinant human tissue TG2 (Sardy *et al*, 1999) that allows early identification of celiac patients.

Hypertrophic scarring occurs in a significant number of patients following surgery or serious burns. A higher percentage of type III collagen due to elevated levels of TG activity is associated with these scars. The use of putrescine as a TG inhibitor has been patented for treatment of the scar tissue (Dolynchuk and Bowness, 1999).

This patent led to a topical cream which has been proven to significantly reduce hypertrophic scarring in clinical trials.

TGs have been implicated in other pathological processes such as tumour growth and metastasis, arteriosclerosis, neurodegenerative diseases and skin pathology (Griffin *et al*, 2002). Clarification of the role of TG in these processes is a very active area of research and there is strong evidence that TG enzymes or their inhibitors offer big potential as therapeutic agents for these processes, and in some cases medical applications have already been patented (Steinert *et al*, 2000).

5.7. Cosmetic Industry

The important roles that TGs play in epidermal tissue differentiation determines that modulation of the activity of these enzymes will affect the structure of the skin and can be used to improve skin appearance. Topical preparations consisting of TG and one or more of the corneocyte proteins are proposed to form a protective layer on the surface of hair, skin and nails (Green and Dijan, 1996). It has been suggested that appearance of wrinkles associated with ageing can be retarded by the potentiation of TG1 expression by the topical application of plant extracts (Megata, 2004). In addition, TG can be used to covalently bind various primary amine-containing compounds, such as antimicrobials, UV-absorbers, anti-inflammatory substances, antioxidants, colouring agents, perfumes and insect repellants to skin, nails and hair surfaces (Richardson *et al*, 1996). A method has been describe for curl retention in hair and lashes based on the action of the enzyme TG (Mammone and Popescu, 2004).

5.8. Textile and Leather

Treatment of leather with TG, preferably together with a glutamine and/or lysine containing polymer such as keratin or casein, has a beneficial effect on the subsequent dyeing and colour properties of leather (Collighan *et al*, 2002). Transglutaminase treatment is also shown to improve resistance to abrasion and collagenase activity (Addy *et al*, 2005).

The treatment of wool textiles with TG has been shown to improve the qualities of finished garments: shrink resistance, handle, wettability, reduction of felting tendency, improved softness, tensile strength retention, improved stretch, improved dye uptake, dye wash-fastness, and protection from proteolytic detergents (McDevitt and Winkler, 1999; Cortez *et al*, 2005). Using a similar methodology to that applied to human skin, hair and nail treatments, TG can be used to covalently bind primary amine-containing active compounds or proteins to fibrous textiles such as animal wool (Griffin *et al*, 2003).

5.9. Other applications

Many authors have described applications of this enzyme as an enzymatic cross-linking agent useful for the generation of analytical tools for immunoassays or

biosensors. In this context, it can be used to conjugate proteins (Tanaka *et al*., 2004) and create hapten-antibody conjugates (Meusel, 2004), labelling of proteins (Josten *et al*., 2000), or to immobilize proteins onto solid surfaces (Kamiya *et al*., 2005).

REFERENCES

Addy, V., Griffin, M., Collighan, R. and Clara, S. (2005) Enzymatic treatment of hides and leather. World patent: WO2005007897.

Aeschlimann, D. and Paulsson, M. (1994) Transglutaminase: Protein cross-linking enzyme in tissues and body fluids. Thromb. Haemost. *71*, 402–415.

Aeschlimann, D., Koeller, M.K., Allen-Hoffmann, B.L. and Mosher, D.F. (1998) Isolation of a cDNA Encoding a Novel Member of the Transglutaminase Gene Family from Human Keratinocytes. Detection and identification of transglutaminase gene products based on reverse transcription-polymerase chain reaction with degenerate primers. J. Biol. Chem. *273*, 3452–3460.

Ahvazi, B., Kim, H.C., Kee, S.-H., Nemes, Z. and Steinert, P.M. (2002) Three-dimensional structure of the human transglutaminase 3 enzyme: binding of calcium ions changes structure for activation. EMBO J. *21*, 2055–2067.

Akimov, S.S., Krylov, D., Fleischman, L.F. and Belkin, A.M. (2000) Tissue transglutaminase is an integrin-binding adhesion coreceptor for fibronectin. J Cell Biol. *148*, 825–838.

Ando, H., Adachi, M., Umeda, K., Matsuura, A., Nonaka, M., Uchio, R., Tanaka, H. and Motoki, M. (1989) Purification and characteristics of a novel transglutaminase derived from micro-organisms. Agric Biol Chem. *53*, 2613–2617.

Candi, E., Schmidt, R. and Melino, G. (2005) The cornified envelope: a model of cell death in the skin. Nat Rev Mol Cell Biol. *6*, 328–340.

Collighan, R., Cortez, J. and Griffin, M. (2002) The biotechnological applications of transglutaminases. Minerva Biotecnologica. *14*, 143–148.

Cortez, J., Bonner, P.L. and Griffin, M. (2005) Transglutaminase treatment of wool fabrics leads to resistance to detergent damage. J. Biotechnol. *116*, 379–386.

Coussons, P.J., Price, N.C., Kelly, S.M., Smith, B. and Sawyer, L. (1992) Factors that govern the specificity of transglutaminase-catalysed modification of proteins and peptides. Biochem J. *282*, 929–930.

Del Duca, S. and Serafini-Fracassini, D. (2005) Transglutaminases of higher, lower plants and fungi. Prog Exp Tumor Res. *38*, 223–247.

Della Mea, M., Caparros-Ruiz, D., Claparols, I., Serafini-Fracassini, D. and Rigau, J. (2004) AtPng1p. The first plant transglutaminase. Plant Physiol. *135*, 2046–2054.

Della Mea, M., Di Sandro, A., Dondini, L., Del Duca, S., Vantini, F., Bergamini, C., Bassi, R. and Serafini-Fracassini, D. (2004) A *Zea mays* 39-kDa thylakoid transglutaminase catalyses the modification by polyamines of light-harvesting complex II in a light-dependent way. Planta. *219*, 754–764.

Dickinson, E. and Yamamoto, Y. (1996) Rheology of milk protein gels and protein-stabilized emulsion gels cross-linked with transglutaminase. J. Agr Food Chem. *44*, 1371–1377.

Dieterich, W., Ehnis, T., Bauer, M., Donner, P., Volta, U., Riecken, E. and Schuppan, D. (1997) Identification of tissue transglutaminase as the autoantigen of celiac disease. Nat Med. *3*, 797–801.

Dolynchuk, K.N. and Bowness, J.M. (1999) Use of transglutaminase inhibitor for the treatment of scar tissue. US Patent: US5885982.

Folk, J.E. (1969) Mechanism of action of guinea pig liver transglutaminase. VI. Order of substrate addition. J Biol Chem. *244*, 3707–3713.

Folk, J.E. and Chung, S.I. (1973) Molecular and catalytic properties of transglutaminases. Advances in Enzymology. *38*, 109–191.

Fox, B.A., Yee, V.C., Pedersen, L.C., Le Trong, I., Bishop, P.D., Stenkamp, R.E. and Teller, D.C. (1999) Identification of the Calcium Binding Site and a Novel Ytterbium Site in Blood Coagulation Factor XIII by X-ray Crystallography. J. Biol. Chem. *274*, 4917–4923.

Gerrard, J.A., Fayle, S.E., Wilson, A.J., Newberry, M.P., Ross, M. and Kavale, S. (1998) Dough properties and crumb strength of white pan bread as affected by microbial transglutaminase. J. Food Sci. *63*, 472–475.

Gerrard, J.A., Newberry, M.P., Ross, M., Wilson, A.J., Fayle, S.E. and Kavale, S. (2000) Pastry lift and croissant volume as affected by microbial transglutaminase. J. Food Sci. *65*, 312–314.

Green, H. and Dijan, P. (1996) Cosmetic containing corneocyte proteins and transglutaminase, and method of application. US Patent: US5525336.

Grenard, P., Bates, M.K. and Aeschlimann, D. (2001) Evolution of transglutaminase genes: identification of a transglutaminase gene cluster on human chromosome 15q15. Structure of the gene encoding transglutaminase X and a novel gene family member, transglutaminase Z. J Biol Chem. *276*, 33066–33078.

Griffin, M., Casadio, R. and Bergamini, C.M. (2002) Transglutaminases: nature's biological glues. Biochem J. *368*, 377–396.

Griffin, M., Cortez, J.M. and Bonnes, P. (2003) Method for enzymatic treatment of textiles such as wool. US Patent: US2003154555.

Haejung, A., Peters, M.Y. and Seymour, T.A. (1996) Roles of endogenous enzymes in surimi gelation. Trends Food Sci and Tech. *7*, 321–327.

Horiguchi, Y. (2001) Escherichia coli cytotoxic necrotizing factors and Bordetella dermonecrotic toxin: the dermonecrosis-inducing toxins activating Rho small GTPases. Toxicon. *39*, 1619–1627.

Huang, Y.P., Seguro, K., Motoki, M. and Tawada, K. (1992) Cross-linking of contractile proteins from skeletal muscle by treatment with microbial transglutaminase. J Biochem (Tokyo). *112*, 229–234.

Ichinose, A., Bottenus, R.E. and Davie, E.W. (1990) Structure of transglutaminases. J Biol Chem. *265*, 13411–13414.

Josten, A., Haalck, L., Spener, F. and Meusel, M. (2000) Use of microbial transglutaminase for the enzymatic biotinylation of antibodies. J Immunol Methods. *240*, 47–54.

Kamiya, N., Doi, S., Tominaga, J., Ichinose, H. and Goto, M. (2005) Transglutaminase-mediated protein immobilization to casein nanolayers created on a plastic surface. Biomacromolecules. *6*, 35–38.

Kanaji, T., Ozaki, H., Takao, T., Kawajiri, H., Ide, H., Motoki, M. and Shimonishi, Y. (1993) Primary structure of microbial transglutaminase from Streptoverticillium sp. strain s-8112. J Biol Chem. *268*, 11565–11572.

Kashiwagi, T., Yokoyama, K., Ishikawa, K., Ono, K., Ejima, D., Matsui, H. and Suzuki, E. (2002) Crystal Structure of Microbial Transglutaminase from Streptoverticillium mobaraense. J. Biol. Chem. *277*, 44252–44260.

Kobayashi K, Kumazawa, Y., Miwa, K. and Yamanaka, S. (1996) epsilon-(gamma-glutamyl)lysine cross-links of spore coat proteins and transglutaminase activity in Bacillus subtilis. FEMS Microbiol. Let. *144*, 157–160.

Kuraishi, C., Yamazaki, K. and Susa, Y. (2001) Transglutaminase: its utilization in the food industry. Food Rev. Internat. *17*, 221–246.

Larre, C., Denery Papini, S., Popineau, Y., Deshayes, C., Desserme, C. and Lefebvre, J. (2000) Biochemical analysis and rheological properties of gluten modified by transglutaminase. Cereal Chem. *77*, 121–127.

Liu, S., Cerione, R.A. and Clardy, J. (2002) Structural basis for the guanine nucleotide-binding activity of tissue transglutaminase and its regulation of transamidation activity. Proc Natl Acad Sci U S A. *99*, 2743–2747.

Lorand, L. and Conrad, S. (1984) Transglutaminases. Mol. Cell. Biochem. *58*, 9–35.

Lorenzen, P.C., Schlimme, E. and Roos, N. (1998) Crosslinking of sodium caseinate by a microbial transglutaminase. Nahrung. *42*, 151–154.

Mammone, T. and Popescu, L. (2004) Method for curl retention in hair and lashes. World Patent Application WO 2004/052279A3.

Mangala, L.S., Arun, B., Sahin, A.A. and Mehta, K. (2005) Tissue transglutaminase-induced alterations in extracellular matrix inhibit tumor invasion. Mol Cancer. *4*, 33.

Matheis, G. and Whitaker, J.R. (1987) A review: enzymatic cross-linking of proteins applicable to foods. J Food Biochem. *11*, 309–327.

Matsumura, Y., Chanyongvorakul, Y., Kumazawa, Y., Ohtsuka, T. and Mori, T. (1996) Enhanced susceptibility to transglutaminase reaction of alpha- lactalbumin in the molten globule state. Biochim Biophys Acta. *1292*, 69–76.

McDevitt, J.P. and Winkler, J. (1999) A method of treating wool, wool fibers or animal hair with a proteolytic enzyme and a transglutaminase. World patent: WO9960200.

Megata, H. (2004) Epidermis keratinization normalizing agent and skin care preparation for external use for skin containing the same. Japan patent: JP2004091376.

Meusel, M. (2004) Synthesis of hapten-protein conjugates using microbial transglutaminase. Methods Mol Biol. *283*, 109–123.

Motoki, M. and Seguro, K. (1998) Transglutaminase and its use for food processing. Trends Food Sci and Tech. *9*, 204–210.

Nemes, Z., Marekov, L.N., Fesus, L. and Steinert, P.M. (1999) A novel function for transglutaminase 1: Attachment of long-chain omega -hydroxyceramides to involucrin by ester bond formation. Proc Natl Acad Sci U S A. *96*, 8402–8407.

Nemes, Z. and Steinert, P.M. (1999) Bricks and mortar of the epidermal barrier. Exp Mol Med. *31*, 5–19.

Nielsen, P. (1995) Reactions and potential industrial applications of transglutaminase. Review of literature and patents. Food Biotech. *9*, 119–156.

Noguchi, K., Ishikawa, K., Yokoyama, K., Ohtsuka, T., Nio, N. and Suzuki, E. (2001) Crystal Structure of Red Sea Bream Transglutaminase. J. Biol. Chem. *276*, 12055–12059.

Nonaka, M., Toiguchi, S., Sakamoto, H., Kawajiri, H., Soeda, T. and Motoki, M. (1994) Changes caused by microbial transglutaminase on physical properties of thermally induced soy protein gels. Food Hydrocolloids. *8*, 1–8.

Osullivan, M., Kelly, A. and Fox, P. (2002) Effect of transglutaminase on the heat stability of milk: a possible mechanism. J Dairy Sci. *85*, 1–7.

Pasternack, R., Dorsch, S., Otterbach, J.T., Robenek, I.R., Wolf, S. and Fuchsbauer, H.-L. (1998) Bacterial pro-transglutaminase from Streptoverticillium mobaraense. Purification, characterisation and sequence of the zymogen. Eur J Biochem. *257*, 570–576.

Pedersen, L.C., Yee, V.C., Bishop, P.D., Le-Trong, I., Teller, D.C. and Stenkamp, R.E. (1994) Transglutaminase factor XIII uses proteinase-like catalytic triad to crosslink macromolecules. Protein Sci. *3*, 1131–1135.

Pedersen, M.H., Hansen, T.K., Sten, E., Seguro, K., Ohtsuka, T., Morita, A., Bindslev-Jensen, C. and Poulsen, L.K. (2004) Evaluation of the potential allergenicity of the enzyme microbial transglutaminase using the 2001 FAO/WHO Decision Tree. Mol Nutr Food Res. *48*, 434–440.

Piacentini, M., Rodolfo, C., Farrace, M. and Autuori, F. (2000) "Tissue" transglutaminase in animal development. Int J Dev Biol. *44*, 655–662.

Pisano, J., Finlayson, J. and Peyton, M. (1968) Cross-link in fibrin polymerized by factor 13: epsilon-(gamma-glutamyl)lysine. Science. *160*, 892–893.

Ragkousi, K. and Setlow, P. (2004) Transglutaminase-mediated cross-linking of GerQ in the coats of Bacillus subtilis spores. J Bacteriol. *186*, 5567–5575.

Richardson, N.K., Schilling, K.M., Polcalyko, D.J. and Bailey, P.L. (1996) Covalent bonding of active agents to skin, hair or nails. US Patent: US5490980.

Sakamoto, H., Kumazawa, Y., Toiguchi, S., Seguro, K., Soeda, T. and Motoki, M. (1995) Gel strength enhancement by addition of microbial transglutaminase during onshore surimi manufacture. J. Food Sci. *60*, 300–304.

Sarang, Z., Molnar, P., Nemeth, T., Gomba, S., Kardon, T., Melino, G., Cotecchia, S., Fesus, L. and Szondy, Z. (2005) Tissue transglutaminase (TG2) acting as G protein protects hepatocytes against Fas-mediated cell death in mice. Hepatology. *42*, 578–587.

Sardy, M., Odenthal, U., Karpati, S., Paulsson, M. and Smyth, N. (1999) Recombinant human tissue transglutaminase ELISA for the diagnosis of gluten-sensitive enteropathy. Clin Chem. *45*, 2142–2149.

Seguro, K., Kumazawa, Y., Ohtsuka, T., Ide, H., Nio, N., Motoki, M. and Kubota, K. (1995) epsilon-(gamma-glutamyl)lysine: hydrolysis by gamma-glutamyltransferase of different origins, when free or protein bound. J. Agr Food Chem. *43*, 1977–1981.

Seguro, K., Kumazawa, Y., Kuraishi, C., Sakamoto, H. and Motoki, M. (1996) The epsilon-(gamma-glutamyl)lysine moiety in crosslinked casein is an available source of lysine for rats. J Nutr. *126*, 2557–2562.

Serafini-Fracassini, D., Del Duca, S. and Beninati, S. (1995) Plant transglutaminases. Phytochemistry. *40*, 355–365.

Steinert, P., Marekov, L. and Nemes, Z. (2000) Methods for ameliorating icthyosiform skin diseases. World patent: WO0078937.

Tanaka, T., Kamiya, N. and Nagamune, T. (2004) Peptidyl linkers for protein heterodimerization catalyzed by microbial transglutaminase. Bioconjug Chem. *15*, 491–497.

Villalobos, E., Santos, M., Talavera, D., Rodriguez-Falcon, M. and Torne, J.M. (2004) Molecular cloning and characterization of a maize transglutaminase complementary DNA. Gene. *336*, 93–104.

Yee, V., Pedersen, L., Le Trong, I., Bishop, P., Stenkamp, R. and Teller, D. (1994) Three-dimensional structure of a transglutaminase: human blood coagulation factor XIII. Proc Natl Acad Sci U S A. *91*, 7296–7300.

Yee, V.C., Pedersen, L.C., Bishop, P.D., Stenkamp, R.E. and Teller, D.C. (1995) Structural evidence that the activation peptide is not released upon thrombin cleavage of factor XIII. Thrombosis Research. *78*, 389–397.

Yokoyama, K., Nio, N. and Kikuchi, Y. (2004) Properties and applications of microbial transglutaminase. Appl Microbiol Biotechnol. *64*, 447–454.

Zotzel, J., Pasternack, R., Pelzer, C., Ziegert, D., Mainusch, M. and Fuchsbauer, H.-L. (2003) Activated transglutaminase from Streptomyces mobaraensis is processed by a tripeptidyl aminopeptidase in the final step. Eur J Biochem. *270*, 4149–4155.

CHAPTER 33

PENICILLIN ACYLASES

DAVID W. SPENCE* AND MARTIN RAMSDEN
GlaxoSmithKline, North Lonsdale Road, Ulverston, United Kingdom
**dave.w.spence@gsk.com*

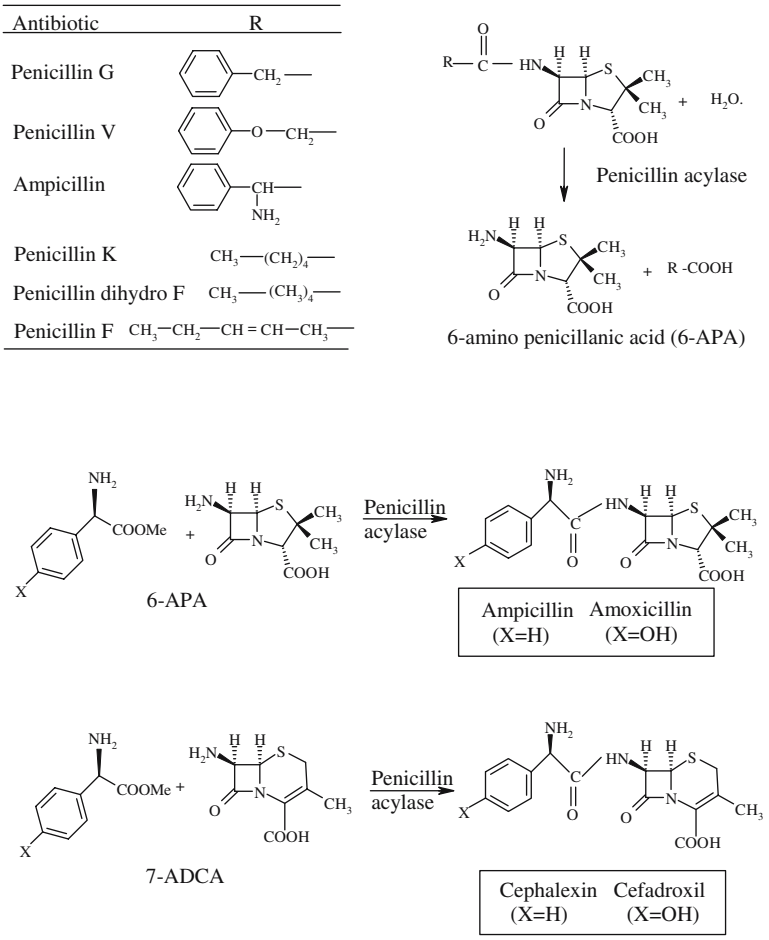
1. INTRODUCTION

The production of penicillin and cephalosporin antibiotics is a multi thousand tonne industrial operation. Benzyl and phenoxymethyl penicillins (penicillin G and penicillin V respectively) are fungal fermentation products and the precursors to a wide range of semi-synthetic antibiotics (amoxicillin, ampicillin *etc*.). The chemical modification of the fermentation product is initiated by removal of the natural acyl group leaving the 6-aminopenicillanic acid (6-APA) penicillin nucleus. Alternative synthetic acyl groups can then be added to confer novel properties to the antibiotic such as resistance to stomach acid, a certain degree of penicillinase resistance or an extended range of antibiotic activity. The method of choice for the conversion to 6-APA at the industrial scale is the use of penicillin acylase. Penicillin acylases or amidases (EC 3.5.1.11) are a group of enzymes which can cleave the acyl chain of penicillins to yield 6-amino penicillanic acid (6-APA) and the corresponding organic acid (Shewale and Sivaraman, 1989), and in a number of cases the same enzyme can be used to direct the synthesis of the new antibiotic by the addition of the novel acyl group (Fig. 1). Apart from their use in modifying β-lactams, these enzymes also have applications in peptide synthesis and the resolution of racemic mixtures of chiral compounds.

2. BIOLOGICAL SOURCES OF PENICILLIN ACYLASES

Penicillin acylases (PAC) are microbial enzymes and are found in a wide range of bacteria, actinomycetes, yeasts and fungi (Suddhakaran and Borkar, 1985 a,b). Based on their substrate specificities (side chain preferences), they have been classified into three groups: penicillin G acylases, penicillin V acylases and ampicillin acylases.

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 583–597



Reactions of Penicillin Acylase

Figure 1. Reactions of Penicillin Acylase

The most abundant, the penicillin G acylases, are found in species including *Escherichia coli* (Cole, 1969), *Alcaligenes faecalis* (Verhaert *et al*, 1997) and *Arthrobacter viscosus* (Ohashi *et al*, 1989). Penicillin V acylases are found in species including *Bacillus subtilis* (Olsson *et al*, 1985) and *Fusarium oxysporium* (Lowe *et al*., 1986), whilst ampicillin acylases are found in *Streptomyces lavendulae* (Torres *et al*., 1999) and *Pseudomonas melangenum* (Kim and Byun, 1990). It should be noted however, that the enzymes from very few of the producing strains are actually used as commercial biocatalysts.

3. ORGANISATION, EXPRESSION AND REGULATION OF THE ACYLASE GENE (*PAC*)

The regulation of penicillin G acylase synthesis has been extensively studied in *E. coli*. This is a good model organism since it is believed that penicillin acylases from diverse sources are processed in a similar manner. There are three major regulatory mechanisms involved in acylase synthesis: induction by phenylacetic acid (PAA), catabolite repression and thermoregulation.

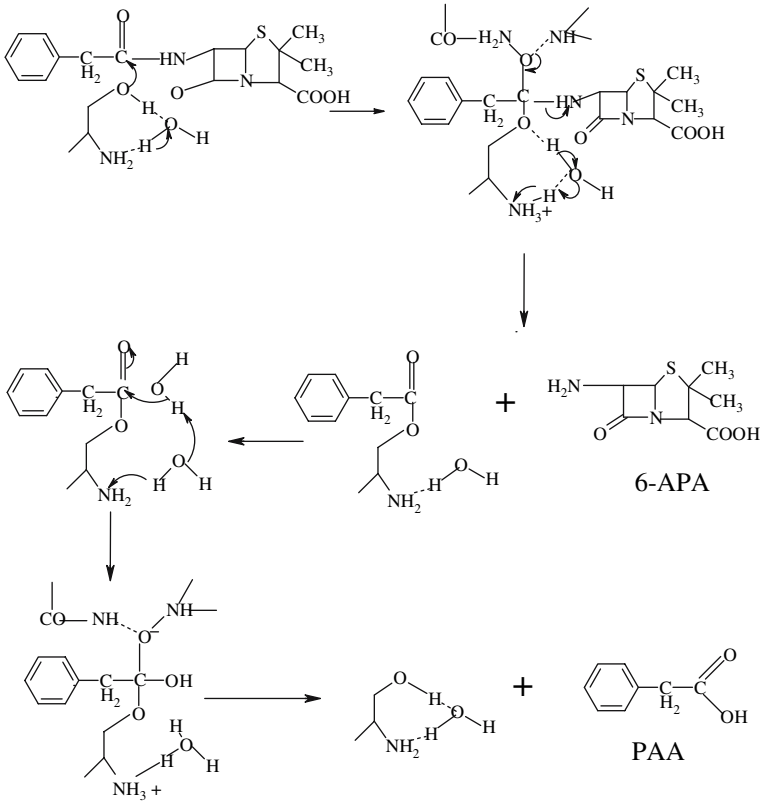
The expression of the wild type *E. coli* acylase gene (*pac*) is induced by PAA and repressed by glucose, these effects being exerted at the transcriptional level (Merino *et al.*, 1992). It had been suggested that a *pac*-specific repressor protein, encoded by a putative *pacR* gene transcribed in the opposite direction to *pac* from within the *pac* coding sequence, is the mediator of the induction (Jiang *et al.*, 1997). However, *lacZ* fusion studies have shown that an intact *pac* gene is not required for PAA induction (Rao and Garcia, 1999). Catabolite repression has been implicated by the presence of putative cAMP receptor protein (CRP)–binding sites upstream of the *pac* gene (Valle *et al.*, 1986). It was also reported that the CRP is directly involved in PAA induction and gene expression (Rao and Garcia, 1999).

PAC is normally produced at temperatures between 20 °C and 30 °C. It is not produced at 37 °C. It is not thought that lack of activity is due to thermal inactivation of the protein since the optimal enzyme activity is 50 °C (Savidge and Cole, 1975). It is more likely that temperature–sensitive proteolytic processing is involved in the maturation of the active enzyme (Oh *et al.*, 1987). In addition, improper folding at the higher temperature results in enzyme inactivity (Lindsay and Pain, 1991). The processing steps involved in the formation of the active peptide subunits are: 1) synthesis of the precursor polypeptide (prepro PAC) which contains a signal peptide, α subunit, spacer peptide and β subunit all encoded by the *pac* gene; 2) transport of the precursor peptide to the periplasmic space and signal peptide removal (pro PAC); 3) proteolytic digestion of the spacer by cleavage at the N terminus of the β peptide; 4) removal of the spacer peptide by C terminal proteolytic processing. A process model has recently been proposed (Ignatova *et al.*, 2005) from *in vitro* folding studies which suggests that an initial fast phase of global burial of hydrophobic residues leads to a stable molten globule–like intermediate state which contains substantial secondary structure with bound calcium ions. This calcium ion, which is found in the mature native enzyme, is responsible for stabilisation rather than contributing directly to protein folding. In the final protein the calcium is found at the junction of the two chains which coordinate it (Kasche *et al.*, 2003). Western blot analysis (Dai *et al.*, 2001) has shown that at elevated temperatures (> 30 °C) the *pac* gene is transcribed and translated but the precursor polypeptide is incorrectly folded resulting in an inactive enzyme. The catabolite repression and PAA induction detailed for *E. coli* can vary in other organisms. For example, in the enterobacteria *Providencia (Proteus) rettgeri* the *pac* gene is not induced by PAA and is repressed by succinate and fumarate rather than glucose (Daumy *et al.*, 1982); in *Alcaligenes faecalis* the gene is induced by PAA but not regulated by

catabolite repression. However, the temperature–dependence of the synthesis of the enzyme appears to be an invariable parameter in all organisms.

4. PROTEIN STRUCTURE AND REACTION MECHANISM

Penicillin acylase is a member of the N terminal nucleophilic (Ntn) hydrolase structural superfamily of proteins, a class of enzymes which share a common fold around the active site and contain a catalytic serine, cysteine or threonine at the N–terminal position (Brannigan *et al*, 1995). PAC has an acyl-binding site which accommodates a PAA group and a nucleophile-binding site which can bind 6-APA or another nucleophile (Fig. 2). It acts through the β chain terminal serine alone to bring about its reaction. The reaction mechanism involves the formation of a covalent intermediate and is similar to that of the serine proteases. After an attack



Proposed mechanism of Penicillin Acylase catalysed hydrolysis of Penicillin G

Figure 2. Proposed mechanism of Penicillin Acylase catalysed hydrolysis of Penicillin G

on the carbonyl carbon of the amide bond by the active site nucleophile, a covalent acyl-enzyme complex is formed via a tetrahedral transition state in which the negatively charged oxyanion is stabilised by H-bonds to oxyanion residues (βN241 and βA69 in the case of *E. coli* PAC) (Duggleby *et al*, 1995). After expulsion of the leaving group from the active site, the acyl-enzyme is deacylated by water, or another nucleophile, giving rise to the final transacylation product and the free enzyme.

PAC can also be used in the catalytically reverse direction to condense an acyl group with a β-lactam nucleus. In this reaction an activated acyl donor, generally an amide or methyl ester of a PAA derivative, acylates the enzyme at the active site serine with the concomitant release of ammonia or methanol. The resulting acyl-enzyme is then deacylated by a β-lactam nucleophile such as 6-APA or 7-desacetoxycephalosporanic acid (7-ADCA). This yields a semisynthetic penicillin or cephalosporin.

Mutation studies coupled with crystal structure data have demonstrated that a number of amino acids within the protein have critical functions in either the processing or catalytic reactions, or both. The βK10 residue is involved in the pH-dependent processing of the preproprotein (Lee *et al*, 2000). Residues involved in catalysis by direct contact include αR145, αF146, βS1, βA127, βN241 and βR263 in *E. coli*, along with the hydrophobic residues αF146, βF24 and βF57 which are implicated in substrate specificity and acyl transfer (Alkema *et al*, 2002). Sequence alignment studies of PAC proteins from sources other than *E. coli* have been performed and in many cases homologous amino acids have been found giving weight to the idea that the mode of action of PAC is the same across diverse species.

The hydrolysis of Pen G to 6-APA proceeds at a slightly alkaline pH, whereas the reverse reaction proceeds at acidic pH. It follows then that the overall reaction is determined by pH and the relative concentrations of substrate and product. Kheirolomoom *et al* (2001) have demonstrated that PAC is inhibited by excess substrate and the presence of both products, 6-APA and PAA. Non-competitive inhibition by 6-APA and competitive inhibition by PAA were observed for the ordered uni bi deacylation reaction in the forward direction. In the reverse direction, Pen G acts as a mixed type inhibitor with respect to 6-APA at unsaturating concentrations of PAA. This inhibition can be overcome with saturating amounts of PAA. In addition, Pen G is a competitive inhibitor with respect to PAA at all concentrations of 6-APA.

5. ENHANCING THE EXPRESSION OF PAC

Classical mutation and selection procedures have been applied to enhance the production of PAC from many micro-organisms. Erarslan *et al* (1991) performed random mutagenesis induced by NTG or NMS followed by screening for overproducers using a *Serratia marcescens* overlay technique. A four-fold improved *E. coli* ATTC 11105 mutant was obtained in this way. Another approach has been the

isolation of mutants in the presence of amides as sole nitrogen sources. The rationale for this is that PAC expression liberates nitrogen for growth. The addition of glucose in the medium has resulted in the isolation of mutants which are resistant to catabolite repression, leading to increased PAC activity (Shewale and Sivaraman, 1989). A succinate resistant mutant of *P. rettgeri* showing increased activity has also been reported (Daumy *et al*, 1985).

Recombinant DNA technology has made it possible to clone and express the *pac* gene from many sources. The gene has been cloned into multicopy vectors to increase gene dosage. This has been achieved from species including *E. coli* (Oh *et al*., 1987), *Arthrobacter viscosus* (Ohashi *et al*, 1989), *P. rettgeri* (Daumy *et al*, 1986), *Kluyvera citrophila* (Barbero *et al*, 1986), *Bacillus megaterium* (Meevootisom and Saunders, 1987), *A. faecalis* (Verhaert *et al*, 1997), and *Achromobacter xylosoxidans* (Cai *et al*, 2004). There are fewer examples of the cloning of the pen V acylase gene since the pen G acylase gene remains the gene of choice for exploitation. Segregational instability of recombinant plasmids in the absence of selective pressure is widespread. To overcome this effect, the *E. coli pac* gene has been cloned in pYA292, which is stable in strain χ6212 in the absence of challenge. The resultant clone produces higher PAC levels (Vohra *et al*, 2001). The location of the mature PAC protein can determine the efficiency of production. For example Ohashi *et al* (1989) showed that expressing the *A. viscosus pac* gene in *E coli* resulted in PAC production intracellularly, whilst expression of the same gene in *B subtilis* resulted in extracellular production with higher levels of activity. The penicillin acylases from gram negative bacteria accumulate within the periplasmic space, whereas those from gram positive bacteria tend to be secreted. Other organisms such as *Bacillus sp* have been shown to produce the enzyme intracellularly (Rajendhran *et al*, 2002).

The transcription of the *pac* gene in *E. coli* is inefficient due to the non-optimal disposition of bases present in the regulatory sequence. The translation of the message is also relatively poor due to there only being four bases between the ATG start codon and the ribosome binding site. PAC expression has been enhanced by placing the *pac* gene under the control of a strong promoter and optimising the spacing upstream of the translational start (Chou *et al*, 1999). However, one of the drawbacks to this approach was the subsequent production of inclusion bodies which reduced the level of PAC activity recovered. Since active PAC is formed in the periplasmic space of *E. coli*, engineering of strains to improve transport and folding would be expected to be beneficial to PAC production. In this regard, Lin *et al* (2001) have demonstrated that in order to avoid solubility problems it is necessary to achieve balanced protein synthesis flux throughout the gene expression steps (transcription, translation and post translational modification) in addition to improvements in the efficiency of each step. This has led to significant improvement in recombinant PAC production. For the commercial manufacture of PAC, combinations of the aforementioned approaches have been applied to production strains to improve the efficiency of the process.

6. CONDITIONS FOR THE PRODUCTION, EXTRACTION AND IMMOBILIZATION OF PAC

The most widely used organism is *E. coli*, containing either the endogenous gene or a heterologous gene. Fermentations are performed in up to 250m^3 fed batch, stirred fermenters. Seed cultures are grown at 37 °C for a prescribed time, usually a pH change indicates optimal time for transfer, then up to 10% seed can be used to inoculate the production fermenter. After an initial growth period, the temperature is reduced and carbon feeding is initiated. This is the trigger for PAC production. The fermentation is continued until maximum activity is achieved and then the fermentation is harvested.

The level of dissolved oxygen is required to be maintained at 15% or higher to avoid repression of PAC. In addition, since glucose is a known repressor of PAC activity when used as carbon source its rate of supply has to be limited. Alternative carbon sources such as sucrose or glycerol are commonly used. The pH of the culture needs to be maintained in the range 6.8 to 7.1 for optimal PAC production. Likewise, the production temperature has to be maintained below 30 °C. Since the expression of the *pac* gene can be differentially modulated depending on the exact nature of the plasmid constructed, diverse expression inducers can be used to switch on the *pac* gene fused to a particular upstream sequence. Examples include galactose (De Leon *et al*., 2003) and isopropyl-beta-D-thiogalactopyranoside (IPTG) (Wen *et al*., 2005) to induce the *lac* promoter, and rhamnose (Deak *et al*., 2003) to switch on the *rhaBAD* promoter.

Since the active PAC enzyme is localised in the periplasmic space in *E. coli* this immediately affords a purification step as less than 5% of cellular proteins are found in this compartment. The periplasmic fraction can be separated from the cytoplasmic fraction using selective permeabilisation. Cell permeabilisation by osmotic shock in combination with EDTA has been reported to yield 70% protein without affecting cell viability (Neu and Heppel, 1965). Polar organic solvents and aqueous solutions of ionic and non-ionic detergents have been used to permeabilise cells (Novella *et al*., 1994). For example, on a laboratory scale, treatment with guanidine/EDTA has been reported to give a 93% yield with a 25-fold purification (Novella *et al*., 1994. The use of organic solvents as permeabilisation agents has also been demonstrated (De Leon *et al*., 2003). The authors reported attempts with ten different solvents with varying results. Solvents with dielectric constants below five and having high hydrophobicity released the most active PAC protein. The production of reverse micelles using AOT in water/hexane resulted in the selective release of the enzyme without cell breakage (Bansal-Mutalik and Gaikar, 2003). More specific means of purification of PAC have been demonstrated. Sanchez *et al*. (2001) achieved greater than 65–fold purification by immobilized metal affinity chromatography (IMAC). The use of pseudo-affinity chromatography to aid the purification of PAC has been expertly reviewed by Fitton *et al*. (2001). The advantage of this technology is that it can be scaled up and the matrices are suitable for sanitisation.

Whilst the permeabilisation methods outlined above can be applied to small scale cultures, the release of PAC from industrial fermentations poses greater problems.

To this end the large scale release of PAC usually involves physical disruption followed by partial purification. After fermentation cells are harvested by centrifugation or settled with flocculant. The concentrated cells are then homogenised and cell debris removed. The extract is then purified by the method of choice. The most general means of purification is chromatography and/or ammonium sulphate precipitation. The exact method depends on the given specific activity and the activity required for industrial use as a biocatalyst. A typical method may involve a pH shift and heating to selectively denature sensitive proteins leaving PAC active in the mix.

The use of two-phase affinity partitioning as a means of protein purification is well documented. A polyethyleneglycol (PEG) derivative and salt system has been shown to be effective in the purification of PAC (Gavasane and Gaikar, 2003). The PEG derivative is selected on the basis of its expected interaction through hydrophobic, electrostatic and biospecific effects. In most cases the ligand has a structure analogous to the penicillin G substrate of the enzyme. Thus benzoate and phenylacetamide derivatives are useful ligands. The salt phase usually contains sodium sulphate or sodium citrate (Marcos *et al*, 1999).

Thermostability can be an important feature for enzymes used in industrial processes as higher temperatures can be used to enhance reaction rates, shift thermodynamic equilibria, increase reactant solubility and decrease the reaction viscosity (Kazan and Erarslan, 1997). The catalytic performance of PAC increases at temperatures between 25 °C and 50 °C. However, the enzyme shows poor stability at temperatures above 35 °C and therefore an effective means of providing thermal stability is highly desirable. Many methods of PAC thermostabilisation have been studied. Thermal unfolding of penicillin acylases has been linked to their conformational mobility in water. This mobility can be reduced by diminishing the amount of free water which can be achieved by the addition of stabilisers such as polyol compounds (Erarslan, 1995), bisimidoesters (Erarslan and Ertan, 1995), neutral salts or proteins. The stability of PAC was improved by up to 180% by the addition of trehalose after thirty minutes incubation at 60 °C (Azevedo *et al*, 1999).

Another approach to attaining thermostability is to isolate inherently thermostable PAC proteins from natural sources. Examples of this include the cloning and expression in *E. coli* of the *pac* gene from *A. faecalis* (Verhaert *et al*, 1997) and *Ac. xylosoxidans* (Cai *et al*, 2004).

In order to be useful as a biocatalyst the PAC enzyme preparation has to be active, robust and for economic reasons, re-usable. One of the most effective ways to enhance stability for many enzymes is to immobilise the enzyme onto a solid support. In addition, immobilisation may allow re-use of the catalyst and thus increase cost effectiveness. A number of immobilisation methods have been used in this context. Each method shows superiority to the more traditional use of free cells, extracts or even immobilized whole cells. Immobilized enzyme preparations attain higher activity and specificity and show better control of contamination. The methods used include adsorption, fibre entrapment, microencapsulation, cross-linking, copolymerisation and covalent attachment (reviewed by Shewale and Sivaraman, 1989).

The most common immobilisation methods include cross-linking and covalent attachment. Glutaraldehyde is the usual cross-linking agent used. A 15-fold improvement in thermostability was apparent after cross-linking with dimethyladipimate (Erarslan and Ertan, 1995). The use of different physical forms of chitosan (powder, particles or beads) to immobilise PAC either by adsorption followed by reticulation with glutaraldehyde or by direct cross-linking to the matrix pretreated with glutaraldehyde has been reported (Braun *et al*, 1989). Singh *et al* (1988) used double entrapment methodology for the immobilisation of PAC on agar-polyacrylamide resins. A highest specific activity of 322U/g was obtained by covalent binding of PAC onto vinyl copolymers (Dahl *et al*, 1985). At present, commercial manufacturers such as Resindion supply epoxy group resins such as Sepabeads™ for use as immobilisation matrices for enzymes including PAC. Sepabeads are porous spherical beads with outstanding mechanical stability and extensive cross-linking. Multipoint covalent immobilisation of PAC from *K. citrophila* stabilised the enzyme 10,000-fold compared to the soluble enzyme from *E. coli* (Guisan *et al*, 1993).

A further means of PAC enzyme stabilisation is afforded by the production of cross-linked enzyme crystals or CLECs (SynthaCLEC-PA, Altus Biologics). This approach is unique in that it results in both stabilisation and immobilisation without activity dilution. The protein matrix is both the catalyst and the support. The crystals are produced by stepwise crystallisation of the purified enzyme followed by molecular cross-linking to preserve the structure, resulting in a biocatalyst which is extremely stable to both temperature and organic solvents. The stabilisation is a consequence of both polar and hydrophobic interactions. The PAC CLEC has been commercialised for both hydrolytic and synthetic activity use, and retains activity over more than 1000 batches (Govardhan, 1999). Another novel preparation called a cross-linked enzyme aggregate (CLEA) has also been reported (Cao *et al*, 2001). CLEAs are prepared by slowly adding a precipitant such as ammonium sulphate to the enzyme at low temperature. The aggregated enzyme is subsequently linked with glutaraldehyde and is then available for use as a biocatalyst. The CLEA can be used in aqueous media in both forward and reverse reactions. Unlike the CLEC preparation, the CLEA enzyme need not be purified to near homogeneity.

7. INDUSTRIAL USES OF PAC

By far the most widespread use of PACs is in the production of 6-APA from both Pen G and Pen V. Indeed, the annual worldwide production of 6-APA is greater than 10,000 tonnes. Immobilized PAC enzymes mainly from *E. coli, B. megaterium and A. faecalis* are available from a number of commercial suppliers. Reactions are carried out at >5,000L scale under controlled conditions, the pH being either controlled at approximately 8.0, or slowly ramped from 7.0 to, depending upon the catalyst, as high as 8.5. Exposure to high temperature (>30 °C) and pH (>8.0) is minimised to reduce inactivation of the enzyme and retain high product yield of the otherwise relatively unstable 6-APA.

The use of PAC in large scale production of semisynthetic penicillins and cephalosporins is also widespread. These processes are focused on the condensation of an appropriate D-amino acid derivative with a β-lactam nucleus in a PAC-catalysed reaction (see Wegman *et al*, 2001). This involves the direct acylation of nucleophiles such as 6-APA or 7-ADCA with free acids at low (<7.0) pH.

Kinetically controlled synthesis involves an acyl group transfer reaction in which activated acids, esters or amides are used as the acylating agents. The yield of this type of reaction is dependant upon three different reactions carried out by the enzyme: 1) the synthesis of the β-lactam compound, 2) the hydrolysis of the activated acyl donor, and 3) the hydrolysis of the product. There are many ways to optimise such a reaction including optimising pH (Ospina *et al*, 1996), addition of suitable solvents (Rosell *et al*, 1998) and the use of high concentrations of acyl donor and nucleus (Kaasgard and Veitland, 1992). Using such controlled strategies a number of different antibiotics are produced including amongst others the high volume antibiotic products amoxicillin (Diender *et al*, 2000), ampicillin (Illanes and Fajardo, 2001), cephalexin (Hernandez-Justiz *et al*, 1999) and cefazolin (Park *et al*., 2000).

An alternative use of PAC is in peptide synthesis. The acylase can be used for the protection and deprotection of amino groups of amino acids by direct enzymatic synthesis and acyl group transfer reactions. For example PAC has been used as a biocatalyst in the synthesis of the sweetener aspartame (Fuganti *et al*, 1986), and further use has been in the preparation of D-phenyl dipeptides whose esters readily undergo ring closure to the corresponding diketopiperazines. Such peptides are used as food additives and as synthons for fungicidal, antiviral and anti-allergenic compounds (Van Langen *et al*, 2000). In addition, PAC can hydrolyse phenylacetyl derivatives of a number of peptides and resolve enantiomers of some organic compounds (Shewale *et al*, 1990).

PACs have also been shown to resolve racemic mixtures of chiral compounds including amino acids (Ng *et al*, 2001), β-amino esters (Roche *et al*, 1999), amines (Van Rantwijk *et al*, 2002) and secondary alcohols (Svedas *et al*, 1996) in aqueous media, anhydrous organic media and in water-cosolvent mixtures. The pure enantiomers which result are used as intermediates in the production of bio-active compounds. Specific examples include the resolution of ethyl-3-amino-4-pentynoate to generate the S isomer which is used in the synthesis of the anti-platelet agent Xemilofiban (Topgi *et al*, 1999). Kinetic enantioselective acylation of the racemic azetidine intermediate is catalysed by PAC during the synthesis of the cephalosporin antibiotic Loracarbef, an analogue of Cefaclor (Cainelli *et al*, 1997). A further use of immobilized PAC is in the hydrolysis of racemic iso-propylamide of mandelic acid which is used in the preparation of functionalised cephalosporins such as cefamandole and cefonicid (Rocchietti *et al*, 2002).

In the biocatalytic synthesis of ampicillin and cephalexin the acyl chain donor is (*R*)-phenylglycine amide. The resolution of racemic phenylglycinonitrile is carried out in aqueous media by PAC to give the desired enantiomer (Chilov *et al*, 2003). PAC can also be used in organic media, thus increasing substrate solubility.

Indeed, the production of cephalexin in ethylene glycol can be performed with high substrate concentrations and small amounts of enzyme (Illanes *et al*, 2005). Another example is that of the enantioselective acylation of the L-enantiomers of the methyl esters of phenylglycine and 4-hydroxyphenylglycine (Basso *et al*, 2000). This process leads to the isolation of the pure D-enantiomer which is used in the preparation of β-lactam antibiotics.

8. CONCLUDING REMARKS

Penicillin acylase has been studied and exploited for more than fifty years. The gene encoding it has been cloned and sequenced from a number of species and used to further the understanding not only of prokaryotic expression control but also protein maturation and folding. The crystal structure of the enzyme has been solved and the chemistry of cleavage of its substrate is well understood. Bulk production and use of PAC is an important contributor to the wide availability and low price of many penicillin and cephalosporin antibiotics. Exploitation of this industrial enzyme is continuing in a number of applications, with additional substrates and products, and genetic modifications to enhance both the enzyme's activity and its properties. It is clear that this family of enzymes will continue to play an important role in the pharmaceutical and bioprocess world for some considerable time.

REFERENCES

Alkema, W.B.L., Dijkhuis, A-J., de Vries, E. and Janssen, D.B. (2002). The role of hydrophobic active site residues in substrate specificity and acyl transfer activity of penicillin acylase. Eur. J. Biochem. *269*, 2093–2100.

Azevedo, A.M., Fonseca, L.P. and Prazeres, D.M.F. (1999). Stability and stabilisation of penicillin acylase. J. Chem. Technol. Biotechnol. *74*, 1110–1116.

Bansal-Mutalik, R. and Gaikar, V.G. (2003). Cell permeabilisation for extraction of penicillin acylase from *Escherichia coli* by reverse micellar solutions. Enz. Microbiol. Technol. *32*, 14–26.

Barbero, J.L., Buesa, J.M., Buitrago, G.G., Mendez, E., Perez-Aranda, A. and Garcia, J.L. (1986). Complete nucleotide sequence of the penicillin acylase gene from *Kluyvera citrophila*. Gene *49*, 69–80.

Basso, A., Braiuca, P., De Martin, L., Ebert, C., Gardossi, L. and Linda, P. (2000). D-Phenylglycine and D-4-hydroxyphenylglycine methyl esters via penicillin G acylase catalysed resolution in organic solvents. Tetrahedron Asym. *11*, 1789–1796.

Brannigan, J.A., Dodson, G., Duggleby, H.J., Moody, P.C., Smith, J.L., Tomchick, D.R. and Murzin, A.G. (1995). A protein catalytic framework with an N–terminal nucleophile is capable of self-activation. Nature *378*, 416–419.

Braun, J., LeChanu, P. and Le Goffic, F. (1989). The immobilisation of penicillin G acylase on chitosan. Biotechnol. Bioeng. *33*, 242–246.

Cai, G., Zhu, S., Yang, S., Zhao, G. and Jiang, W. (2004). Cloning, overexpression and characterisation of a novel thermostable penicillin G acylase from *Achromobacter xylosoxidans*: Probing the molecular basis for its high thermostability. Appl. Environ. Microbiol. *70*, 2764–2770.

Cainelli, F., Giacomini, D., Galletti, P. and Cacol, M. (1997). Penicillin G acylase mediated synthesis of the enantiopure (*S*)–3-aminoazetidin-2-one. Tetrahedron Asym. *8*, 3231–3235.

Cao, L., Langen, L.M., Rantwijk, F. and Sheldon, R.A. (2001). Cross-linked aggregates of penicillin acylase: robust catalysts for the synthesis of β-lactam antibiotics. J. Mol. Catal. B: Enzym. *11*, 665–670.

Chilov, G.G., Moody, H.M., Boesten, W.H.J. and Svedas, V.K. (2003). Resolution of (RS)-phenylglycinonitrile by penicillin acylase-catalysed acylation in aqueous medium. Tetahedron Asym. *14*, 2613–2617.

Chou, C.P., Yu, C., Tseng, J., Lin, M. and Lin, H. (1999). Genetic manipulation to identify limiting steps and develop strategies for high level expression of penicillin acylase in *Escherichia coli*. Biotechnol. Bioeng. *63*, 263–272.

Cole, M. (1969). Hydrolysis of penicillins and related compounds by the cell–bound penicillin acylase of *Escherichia coli*. Biochem. J. *115*, 733–739.

Dahl, P., Babu, G.N., Sudhakaran, S. and Borkar, S.P. (1985). Immobilisation of penicillin acylase by covalent linkage on vinyl copolymers containing epoxy groups. Makromol. Chem. Rapid Commun. *6*, 91–95.

Dai, M., Zhu, K., Yang, Y., Wang, E., Xie, Y., Zhao, G. and Jiang, W. (2001). Expression of penicillin G acylase from the cloned *pac* gene of *Escherichia coli* ATCC 11105. Effects of *pacR* and temperature. Eur. J. Biochem. *268*, 1298–1303.

Daumy, G.O., Danley, D., McColl, A.S., Apostolakos, D. and Vinick, F.J. (1985). Experimental evolution of penicillin G acylase from *Escherichia coli* to *Proteus rettgeri*. J. Bacteriol. *163*, 925–932.

Daumy, G.O., McColl, A.S. and Apostolakos, D. (1982). Repression of penicillin G acylase of *Proteus rettgeri* by tricarboxylic acid intermediates. J. Bacteriol. *152*, 104–110.

Daumy, G.O., Williams, J.A., McColl, A.S., Zuzel, T.J. and Danley, D. (1986). Expression and regulation of the penicillin G acylase gene from *Proteus rettgeri* cloned in *Escherichia coli*. J.Bacteriol. *168*, 431–433.

Deak, P.M., Lutz-Wahl, S., Bothe, H. and Fischer, L. (2003). Bioreactor cultivation of *Escherichia coli* for production of recombinant penicillin G amidase from *Alcaligenes faecalis*. Biotechnol. Letts. *25*, 397–400.

De Leon, A., Breceda, G.B., Barba de la Rosa, A.P., Jimenez-Bremont, J.F. and Lopez-Revilla, R. (2003). Galactose induces the expression of penicillin acylase under control of the *lac* promoter in recombinant *Escherichia coli*. Biotechnol. Letts. *25*, 1397–1402.

De Leon, A., Garcia, B., Barba de la Rosa, A.P., Villasenor, F., Estrada, A. and Lopez-Revilla, R. (2003). Periplasmic penicillin G acylase activity in recombinant *Escherichia coli* cells permeabilised with organic solvents. Proc. Biochem. *39*, 301–305.

Diender, M.B., Straathoff, A.J.J., van der Does, T., Zomerdijk, M. and Heijnen, J.J. (2000). Course of pH in the formation of amoxicillin by a suspension-to-suspension reaction. Enz. Microb. Technol. *27*, 576–582.

Duggleby, H.J., Tolley, S.P., Hill, C.P., Dodson, E.J., Dodson, G. and Moody, PC. (1995). Penicillin acylase has a single amino acid catalytic centre. Nature *373*, 264–268.

Erarslan, A. (1995). The effect of polyol compounds on the thermostability of PGA from a mutant strain of *E coli* ATCC11105. Process Biochem. *30*, 133–139.

Erarslan, A. and Ertan, H. (1995). Thermostabilisation of penicillin G acylase obtained from a mutant strain of *E coli* ATCC11105 by bisimidoesters as homobifunctional cross-linking agents. Enz. Microbiol. Technol. *17*, 629–635.

Erarslan, A., Terzi, I., Guray, A. and Bermek, E. (1991). Purification and kinetics of penicillin G acylase from a mutant strain of *Escherichia coli* ATTC 11105. J. Chem. Technol. Biotechnol. *51*, 27–40.

Fitton, V., Verdon, N., Sanchez, J. and Santarelli, X. (2001). Penicillin acylase purification with the aid of pseudo-affinity chromatography. J. Biochem. Biophys. Methods. *49*, 553–560.

Fuganti, C., Grasselli, P. and Casati, P. (1986). immobilized penicillin acylase: application to the synthesis of the dipeptide aspartame. Tetrahedron Letters *27*, 3191–3194.

Gavasane, M.R. and Gaikar, V.G. (2003). Aqueous two-phase affinity partitioning of penicillin acylase from *E coli* in presence of PEG–derivatives. Enz. Microb. Technol. *32*, 665 –675.

Govardhan, C. (1999). Crosslinking of enzymes for improved stability and performance. Curr. Opin. Biotechnol. *10*, 331–335.

Guisan, J.M., Alvaro, G., Fernandez-Lafuente, R., Rosell, C.M., Garcia, J.L. and Tagliani, A. (1993). Stabilisation of heterodimeric enzyme by multipoint covalent immobilisation: PGA from *K. citrophila*. Biotechnol. Bioeng. *42*, 455–464.

Hernandez-Justiz, O., Terrini, M., Pagani, G., Garcia J.L., Guisan, J.M. and Fernandez-Lafuente, R. (1999). Evaluation of different enzymes as catalysts for the production of β-lactam antibiotics following a kinetically controlled strategy. Enz. Microb. Technol. *25*, 336–343.

Ignatova. Z., Wischnewski. F., Notbohm. H. and Kasche. V. (2005). Pro–sequence and Ca^{2+}-binding: Implications for folding and maturation on Ntn-hydrolase penicillin amidase from *E. coli*. J. Mol. Biol. *348*, 999–1014.

Illanes, A., Altamirano, C., Fuentes, M., Zamorano, F. and Aguirre, C. (2005). Synthesis of cephalexin in organic medium at high substrate concentrations and low enzyme to substrate ratio. J. Mol. Cat. B: Enzymatic *35*, 45–51.

Illanes, A. and Fajardo, A. (2001). Kinetically controlled synthesis of ampicillin with immobilized penicillin acylase in the presence of organic cosolvents. J. Mol. Catal. B: Enzymatic *11*, 587–595.

Jiang, L., Wang, Y. and Yang, S. (1997). Location of the regulatory gene in penicillin acylase operon (pacR) of *E coli* D816. Genet. Anal. Biomol. Eng. *14*, 51–54.

Kaasgaard, S. and Veitland, U. (1992). Process for the preparation of β-lactams. International Patent Application WO 92/01061 to Novo-Nordisk.

Kasche, V., Galunsky, B. and Ignatova, Z. (2003). Fragments of pro-peptide activate mature penicillin amidase of *Alcaligenes faecalis*. Eur. J. Biochem. *270*, 4721–4728.

Kazan, D. and Erarslan, A. (1997). Stabilisation of *Escherichia coli* penicillin G acylase by polyethylene glycol against thermal inactivation. Appl. Biochem. Biotechnol. *62*, 1–13.

Kheirolomoom, A., Ardjmand, M., Fazelina, H. and Zakeri, A. (2001). Clarification of penicillin G acylase reaction mechanism. Proc. Biochem. *36*, 1095–1101.

Kim, D.J. and Byun, S.M. (1990). Purification and properties of ampicillin acylase from *Pseudomonas melanogenum*. Biochimica et Biophysica Acta – Protein Structure and Molecular Enzymology *1040*, 12–18.

Lee, H., Park, O.K. and Kang, H.S. (2000). Identification of a new active site for autocatalytic processing of penicillin precursor in *Escherichia coli* ATCC 11105. Biochem. Biophys. Res. Comm. *272*, 199–204.

Lin, W., Kuo, B. and Chou, C.P. (2001). A biochemical engineering approach for enhancing production of recombinant penicillin acylase in *Escherichia coli*. Bioproc. Biosyst. Eng. *24*, 239–247.

Lindsay, C.D. and Pain, R.H. (1991). Refolding and assembly of penicillin acylase, an enzyme composed of two polypeptide chains that result from proteolytic activation. Biochemistry *30*, 9034–9040.

Lowe, D.A., Romancik, G. and Elander, R.P. (1986). Enzymatic hydrolysis of penicillin V to 6–APA by *Fusarium oxysporum*. Biotechnol. Letts. *8*, 151–156.

Marcos, J.C., Fonseca, L.P., Ramalho, M.T. and Cabralm, J.M.S. (1999). Partial purification of penicillin acylase from *Escherichia coli* in polyethylene glycol–sodium citrate aqueous two–phase systems. J. Chromat. B. *734*, 15–22.

Meevootisom, V. and Saunders, J.R. (1987). Cloning and expression of penicillin acylase genes from overproducing strains of *Escherichia coli* and *Bacillus megaterium*. Appl. Microbiol. Biotechnol. *25*, 372–373.

Merino, E., Balbas, P., Recillas, F., Beccerril, B., Valle, F. and Bolivar, F. (1992). Carbon regulation and the role in nature of the *Escherichia coli* penicillin acylase (pac) gene. Mol. Microbiol *6*, 2175–2182.

Neu, H.C. and Heppel, L.A. (1965). The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. J. Biol. Chem. *240*, 3685–3692.

Ng, J.S., Wang P.T., Landis, B.H., Yonan, E.E. and Topgi, R.S. (2001). Method and apparatus for preparation of chiral β-amino acids using penicillin G acylase. International Patent Application US6214609 to Monsanto.

Novella, S., Fargues, C. and Grevillot, G. (1994). Improvement of the extraction of penicillin acylase from *Escherichia coli* by a combined use of chemical methods. Biotechnol. Bioeng. *44*, 379–382.

Oh, S.J., Kim, Y.C., Park, Y.W., Min, S.Y., Kim, I.S. and Kang, H.S. (1987). Complete nucleotide sequence of the penicillin G acylase gene and flanking regions, and its expression in *Escherichia coli*. Gene *56*, 87–97.

Ohashi, H., Katsuta,Y., Nagashima, M., Kamei, T. and Yano, M. (1989). Expression of the *Arthrobacter viscosus* penicillin acylase gene in *Escherichia coli* and *Bacillus subtilis*. Appl. Environ. Microbiol. *55*, 351–356.

Olsson, A., Hagstrom, T., Nilsson, B., Uhlen, M. and Gatenbeck, S. (1985). Molecular cloning of *Bacillus sphaericus* penicillin V acylase gene and its expression in *E. coli* and *B. subtilis.* Appl. Environ. Microbiol. *49*, 1084–1089.

Ospina, S., Barzana, E., Ramirez, O. and Lopez-Munguia, A. (1996). Effect of pH in the synthesis of ampicillin by penicillin acylase. Enz. Microb. Technol. *19*, 462–469.

Park, C.B., Lee, S.B. and Ryu, D.D.Y. (2000). Penicillin acylase catalysed synthesis of cefazolin in water-solvent mixtures: enhancement effect of ethyl acetate and carbon tetrachloride on the synthetic yield. J. Mol. Catal. B: Enzymatic *9*, 275–281.

Rajendhran, J., Krishnakumar, V. and Gunasekaran, P. (2002). Optimisation of a fermentation medium for the production of penicillin G acylase from *Bacillus sp*. Lett. Appl. Mcrobiol. *35*, 523–527.

Rao, A. and Garcia, J.L. (1999). New insights into the regulation of the *pac* gene from *Escherichia coli* W ATCC 11105. FEMS Microbiol. Lett. *177*, 7–14.

Rocchietti, S., San Vicente, A., Pregnolato, M., Tagliani, A., Guisan, J.M., Fernandez-Lafuente, R. and Terrini, M. (2002). Influence of the enzyme derivative preparation and substrate structure on the enantioselectivity of penicillin acylase. Enz. Microb. Technol. *31*, 88–93.

Roche, D., Prasad, K. and Repic, O. (1999). Enantioselective acylation of β-amino esters using penicillin G acylase in organic solvents. Tetrahedron Lett. *40*, 3665–3668.

Rosell, C.M., Terrini, M., Fernandez-Lafuente, R. and Guisan J.M. (1998). A criterion for the selection of monophasic solvents for enzymatic synthesis. Enz. Microb. Technol. *23*, 64–69.

Sanchez, J., Verdoni, N., Fitton, V. and Santarelli, X. (2001). Efficient two–step chromatographic purification of penicillin acylase from clarified *Escherichia coli* ultrasonic homogenate. J. Chromat. B. *753*, 45–50.

Savidge, T.A. and Cole, M. (1975). Penicillin acylase (bacterial). Methods Enzymol. *43*, 705–721.

Shewale, J.G., Deshpande, B.S., Sudhakaran, V.K. and Ambedkar, S.S. (1995). Penicillin acylases: applications and potentials. Process Biochem. *25*, 97–103.

Shewale, J.G. and Sivaraman, H. (1989). Penicillin acylase enzyme production and its applications in the manufacture of 6-APA. Process Biochem. *24*, 146–154.

Singh, D., Goel, R. and Johri, B.N. (1988). Production of 6-aminopenicillanic acid through double entrapped *E. coli* NCIM 2563. Curr. Sci. *57*(22), 1229–1231.

Sudhakaran, V.K. and Borkar, P.S. (1985a). Microbial transformation of β-lactam antibiotics: enzymes from bacteria, sources and study – a summary. Hind. Antibiotic Bull. *27*, 63–119.

Sudhakaran, V.K. and Borkar, P.S. (1985b). Phenoxymethyl penicillin acylases: sources and study – a summary. Hind. Antibiotic Bull. *27*, 44–62.

Svedas, V.K., Savchenko, M.V., Beltser, A.I. and Garunda, F. (1996). Enantioselective penicillin acylase catalysed reactions. Ann. NY Acad. Sci. *799*, 659–669.

Topgi, R.S., Ng, J.S., Landis, B., Wang, P. and Behling, J.R. (1999). Use of enzyme penicillin acylase in selective amidation/amide hydrolysis to resolve 3-amino-4-pentynoate isomers. Bioorg. Med. Chem. *7*, 2221–2229.

Torres, R., Ramon, F., de la Mata, I., Castillon, M.P. and Acebal C. (1999). Enhanced production of penicillin V acylase from *Streptomyces lavendulae*. Appl. Microbiol. Technol. *53*, 81–84.

Valle, F., Gosset, G., Tenorio, B., Oliver, G. and Bolivar, F. (1986). Characterisation of the regulatory region of the *Escherichia coli* penicillin acylase gene. Gene *50*, 119–122.

Van Langen, L.M., Van Rantwijk, F., Svedas, V.K. and Sheldon, R.A. (2000). Penicillin acylase catalysed peptide synthesis: a chemoenzymatic route to stereoisomers of 3,6–diphenylpiperazine-2,5–dione. Tetrahedron Asym. *11*, 1077–1083.

Van Rintwijk, F., Sheldon, R.A., Van Langen, L.M., Khimiouk, A.L., Guranda, D.T. and Svedas, V.K. (2002). Process for the preparation of enantiomerically enriched amines. International Patent Application WO 02/20821 to DSM.

Verhaert, R.D.M., Riemans, A.M., Laan, J., Duin, J. and Quax, W.J. (1997). Molecular cloning and analysis of the gene encoding the thermostable penicillin G acylase from *Alcaligenes faecalis*. Appl. Environ. Microbiol. *63*, 3412–3418.

Vohra, P.K., Sharma, R., Kashyap, D.R. and Tewari, R. (2001). Enhanced production of penicillin G acylase from a recombinant *Escherichia coli*. Biotechnol. Letts. *23*, 531–535.

Wegman, M.A., Janssen, M.H.A., Van Rantwijk, F. and Sheldon R.A. (2001). Towards biocatalytic synthesis of β-lactam antibiotics. Adv. Synth. Catal. *343*, 559–576.

Wen, Y., Feng, M., Yuan, Z. and Zhou, P. (2005). Expression and overproduction of recombinant penicillin G acylase from *Kluyvera citrophila* in *Escherichia coli*. Enz. Microb. Technol. *37*, 233–237.

CHAPTER 34

HYDANTOINASES

YUN-PENG CHAO[1]*, CHUNG-JEN CHIANG[2], JONG-TZER CHERN[1] AND JASON T.C. TZEN[3]

[1]*Department of Chemical Engineering, Feng Chia University*
[2]*Department of Life Science, and*
[3]*Graduate Institute of Biotechnology, National Chung-Hsing University, Taichung, Taiwan, ROC*
**ypchao@fcu.edu.tw*

1. INTRODUCTION

The hydantoin-cleaving enzymes were first discovered in animals and later characterized using partially purified enzymes from animals and plants (Eadie *et al*, 1949; Gaebler and Keltch, 1926). Assigned by the same EC number (EC 3.5.2.2), hydantoinase (HYD) and dihydropyrimidase (DHP) mediate a similar type of hydrolytic reaction but exhibit different substrate specificities and kinetic characteristics. They catalyse the reversible hydrolytic ring opening of the amide bond (-CO-NH-) in five- or six-membered cyclic diamides (Fig. 1). Hydrolysis of the six-membered ring by DHP is the second step in reductive catabolism of pyrimidines in human. This led to the suggestion that HYD may act like the microbial counterpart (Vogels and Van der Drift, 1976). It was later found that the hydrolysis of 5-substituted hydantoins by microbial HYDs produced enantiomerically pure N-carbamoyl amino acids. These intermediates could be further converted to D- or L-amino acids in an irreversible reaction catalysed by the highly stereoselective microbial N-carbamoylases (Olivieri *et al*, 1979). This finding has ushered in a biotechnological interest in the production of optically pure amino acids because of their potential for use in the production of antibiotics, peptide hormones, pyrethroids and pesticides (Owgawa and Shimizu, 2002).

Given the industrial importance of HYDs, natural habitats were intensively screened for the presence of microbes having high levels of HYD activity. By the conventional approach of using hydantoin or its derivatives as sole carbon and nitrogen sources, many microbes with active HYD activity were identified (Yamada *et al*., 1978). This screening revealed HYDs to be widely distributed among different

J. Polaina and A.P. MacCabe (eds.), Industrial Enzymes, 599–606

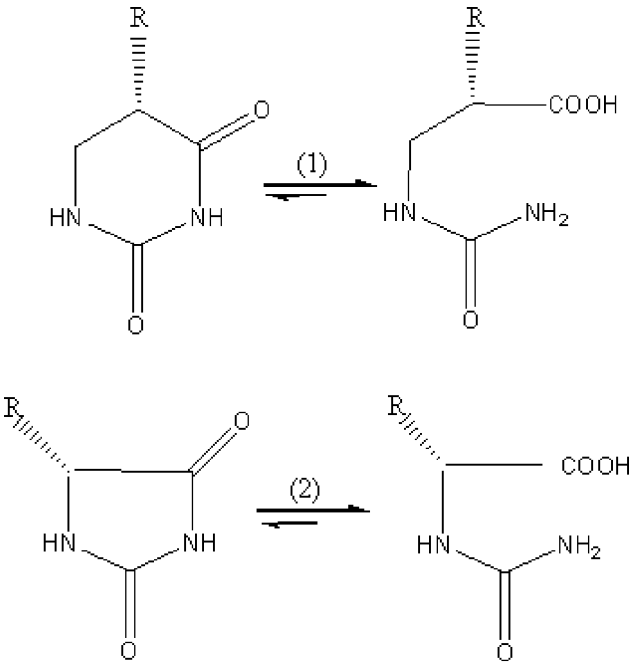


Figure 1. Typical hydrolysis reactions catalysed by (1) dihydropyrimidase and (2) hydantoinase

genera of bacteria. Microbial HYDs display distinct substrate specificities and are generally classified depending on their stereoselectivity. However, this classification is sometimes misleading since it does not provide any information linking the evolutionary relationship of distinct HYDs with their enantioselectivities. For instance, the HYD from *Flavobacterium* sp. was found to be L-selective for indolymethylhydantoin but D-selective for benzyloxymethylhydantoin (Yokozeki *et al*, 1987).

2. BIOLOGICAL FUNCTION AND MOLECULAR STRUCTURE

Whilst a wide variety of HYDs exists in micro-organisms their biological function remains largely unknown. Nevertheless, with the help of computer analysis the amino acid sequences of 16 amidohydrolases (AHL) were aligned and compared (Kim and Kim, 1998). It was found that a long region containing invariant amino acid residues shared 69% similarity among the enzymes analysed. As a result of an extensive survey of the protein data bank, a particular sequence consisting of one aspartic acid and four histidine residues was identified as being highly conserved in the functionally related AHL. Introduction of mutations into this conserved region of the D-HYD gene of *Bacillus stearothermophilus* SD1 (HYDbs) revealed that these invariant residues play a catalytic role. This has ultimately led to the grouping of DHP, allantoinase, dihydroorotase (DHO) and HYD into the

cyclic AHL superfamily due to their similar catalytic functions and rigidity in the conserved sequence (Kim and Kim, 1998).

Our understanding of the catalytic mechanism of HYD has been greatly improved by the recent availability of the crystal structure of HYDbs (Cheon *et al*, 2002). Refined at 3.4 Å resolution, the three-dimensional structure of HYDbs shows that each subunit assumes a classic TIM barrel fold and an extra domain. Structural comparison reveals that HYDbs shares a strikingly resemblance to DHO of *Escherichia coli*. This structural similarity also extends to the active site, thereby suggesting that HYDbs is a di-zinc metalloenzyme: one metal ion activates an attacking hydroxyl while the second metal ion stabilizes the negative charge of the carbonyl oxygen (Thoden *et al*, 2001). Furthermore, the substrate binding pocket of HYDbs was identified by structural suuperimposition of these two enzymes. From a close-up stereodiagram, three stereochemistry gate loops (SGLs) could be clearly discerned. These SGLs were proposed to constitute the hydrophobic core of the substrate binding site and be responsible for the recognition of the exocyclic substituents of hydantoin. This has been evidenced recently by a study that has shown that manipulation of the amino acid residues in the SGLs of HYDbs induces a remarkable change in substrate specificity (Cheon *et al*, 2004). In particular, the enzymatic affinity of mutant HYDbs towards the aromatic substrate hydroxyphenylhydantoin (HPH) could be enhanced 200 fold by the substitution of Phe 159 with Ala. In this context, a working model for the enantioselectivity of HYD is proposed here. As shown in Fig. 2, the HYD substrate binding pocket resembles a barrel with a lid and consists of the recognition sites for the amide group (the barrel bottom) and the functional R group of the substrate (the lid). The functional release of the product from the enzyme occurs only if the chiral portion of the substrate is complementarily coordinated with the active amino acid residues in the hydrophobic core.

A key feature for enzymes of industrial application is their thermostability. Thanks to the information acquired on molecular structures, comparison of a thermostable HYD and HYD from a mesophilic microbe *Burkholderia pickettii*

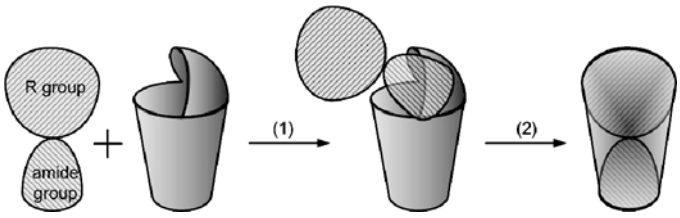


Figure 2. Schematic illustration of a proposed model for the enantioselectivity of hydantoinase. (1) The substrate binding pocket of HYD resembles a barrel with the lid open. (2) The model substrate is guided into the barrel with the amide group downwards. (3) The amide group of the model substrate fits into the barrel bottom (the amide recognition site), while the functional R group is correctly positioned and oriented to make close contact with the lid. As a result, the lid closes and the conformational change leads to the establishment of intimate interactions between the substrate and the protein's active site

(HYDbp) has been made possible (Xu *et al*., 2003), providing an insight into the molecular basis of HYD thermostability. Despite sharing moderate sequence homology, the two HYDs have high similarity as regards their overall and catalytic site structures. The amino acid composition of the substrate recognition site is however less conserved, which accounts for the difference in substrate specificity. Moreover, it indicates that many factors including aromatic residues in the interior of the structure, salt bridges and hydrogen-bonding interactions, and oxygen-sensitive residues on the protein surface likely contribute to the varying thermostability of HYD.

3. PURIFICATION AND IMMOBILIZATION

Progress in the practical application of proteins in industry is usually limited by their production levels and purification processes. Overproduction of microbial HYDs in *E. coli* is prone to the formation of insoluble inclusion bodies. To solve this problem, a systematic study was undertaken to investigate the possible factors affecting the solubility of HYD from *Agrobacterium radiobacter* NRRL B11291 (HYDar) (Chao *et al*., 2000). It was found that an active form of HYDar was produced at low temperature and by coproduction of DnaJ/DnaK. In addition, by fusion with thioredoxin HYDar's solubility could be further improved with the escort of DnaJ/DnaK. Meanwhile, high levels of production of a soluble fusion protein comprising HYDbs and maltose binding protein (MBP) was reported (Kim *et al*., 2000). However, a similar approach constructing a C-terminal fusion of MBP to HYDar yielded the contrary: most of the MBP-fused protein aggregated (unpublished data).

The fusion of affinity tags to a target protein usually makes the work of protein purification easier. The N-terminal fusion of L-HYD of *Arthrobacter aurescens* to MBP permitted a one-step purification of the hybrid protein using an amylase resin (Pietzsch *et al*., 2000). However, the purification of authentic (unmodified) proteins generally requires a tedious procedure consisting of multiple chromatographic steps. To simplify this, we have recently developed a method using a single chromatographic step for the purification of the recombinant HYDar (Huang *et al*., 2003). After the recovery of cell-fee extract, HYDar without any affinity tags was purified to high yield and high purity via a series of steps comprising heat treatment, alcohol precipitation and chelating Sephacel chromatography.

The technique of confining enzymes into a defined space is a very useful approach in many bioprocesses that require the use of enzymes since it facilitates enzyme reutilization without the need for purification. Indeed, enzyme immobilization offers advantages including the easy separation of the latter from the reaction products, the elevation of enzyme concentrations per unit volume, and even the enhancement of enzyme stability and activity. Recently, we have explored the utilization of the chitin binding domain (ChBD) as an affinity tag to retain HYDar on chitin (Chern and Chao, 2005). ChBD of chitinase A1 from *Bacillus circulans* WL-12 comprises 45 amino acids and exhibits remarkably high specificity for chitin (Hashimoto *et al*.,

2000). Compared to its unbound counterpart, immobilized HYDar exhibited higher tolerance to heat and reached a half-life of 270 hr at 45 °C. Notably, HYDar immobilized on chitin could be reused 15 times to achieve conversion yields exceeding 90% using 100-mM HPH as the substrate.

Nevertheless, most of the HYDar produced in *E. coli* remains aggregated though its solubility can be enhanced upon escort by molecular chaperons and the use of low temperatures as mentioned above. Thus, in order to achieve the immobilization of poorly soluble enzymes it is necessary to renature the insoluble proteins, a frequently laborious, inefficient and costly process. A novel method based on artificial oil bodies (AOBs) has recently been proposed (Chiang *et al*., 2006) to achieve protein refolding and immobilization in one step. By fusing HYDar to the C terminus of oleosin, a plant oil body storage protein (Frandsen *et al*., 2001), a hybrid protein was obtained in the form of inclusion bodies in *E. coli* representing 30% of the total cellular protein content. Without the need for addition of a denaturant, the insoluble fusion protein combined with plant oil was readily incorporated into AOBs upon sonication. It was found that in addition to its higher thermal resistance, HYDar immobilized on AOBs could be maintained stable for at least 15 days when stored at 4 °C. When applied repeatedly in a bioconversion process, the immobilized HYDar gave a conversion yield exceeding 80% over 7 cycles. One added advantage of this method is that it allows easy recovery of immobilized enzymes from the top of a solution simply by performing a brief centrifugation. The essence of this novel method for enzyme immobilization is outlined in Fig. 3.

It is well recognized that HYDar mediates the hydrolysis of HPH to N-carbamoyl-D-hydroxyphenyl glycine (CpHPG), an intermediate which can be converted to D-p-hydroxyphenylglcine (D-HPG) (Olivieri *et al*., 1979). D-HPG is of great commercial value for its potential as a precursor of semi-synthetic antibiotics. Thus, the successful achievement of HYD immobilization presents an important and fundamental step in this bioconversion process.

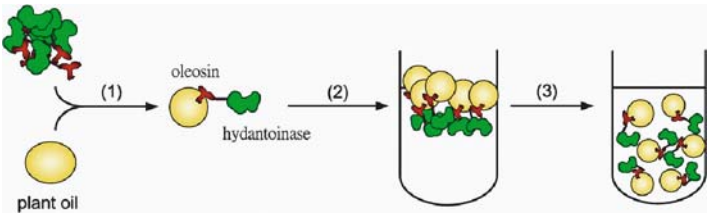


Figure 3. Schematic illustration of the immobilization of hydantoinase on oil bodies. (1) The recovery of the oleosin-fused hydantoinase in the form of inclusion bodies is mixed with plant oil and subjected to sonication. (2) Hydantoinases immobilized on the surface of the oil bodies are removed from the top of the aqueous phase after a brief centrifugation. (3) Oil bodies are fed into the buffer solution to initiate the bioconversion reaction

4. PRODUCTION OF D- OR L-AMINO ACIDS

In principle, enantiomerically pure D- or L-amino acids can be readily obtained from racemic D,L-5-monosubstituted hydantoins in a two-step reaction directed by the combined activity of stereoselective HYD and N-carbamoylase (Fig. 4). This process for the production of optically pure amino acids has received considerable interest by industry. Of particular importance, the production of D-HPG, a precursor used in the synthesis of cephalosporins and obtained from D,L-HPH, was first reported using *A. radiobacter* NRRL B11291 carrying D-HYD and D-carbamoylase (Olivieri *et al*, 1981). To increase the efficiency of this process, a recombinant *E. coli* strain overexpressing these two genes was constructed and used for D-HPG production (Chao *et al*, 1999). A 5-fold increase in D-HPG productivity was achieved by this approach compared to production by *A. radiobacter* NRRL B11291. Moreover, the flux control coefficient of D-carbamoylase for this two-step reaction was determined to be 0.8, suggesting that this enzymatic reaction is the rate-determining step (Chao *et al*, 1999). This analysis suggests the possibility of optimising D-HPG production by modifying the ratio of the two enzyme activities in the reaction. Accordingly, the two respective genes from *A. radiobacter* NRRL B11291 were cloned and expressed separately under the control of the *tac* promoter (Chao *et al*, 2000). With a mixed population consisting of various proportions of recombinant *E. coli* strains producing each individual enzyme, an optimal productivity of D-HPG was obtained with the ratio of D-carbamoylase to D-HYD activity ranging between 1 and 2.

The production of L-amino acids using recombinant *E. coli* strain has also been developed (Wilms *et al*, 2001). Similarly, the balanced activity of L-HYD, L-carbamoylase, and racemase was investigated and established. Application to the reaction process using the engineered *E. coli* cells gave a 6-fold increase in productivity compared to that by the whole cell from *A. aurescens*.

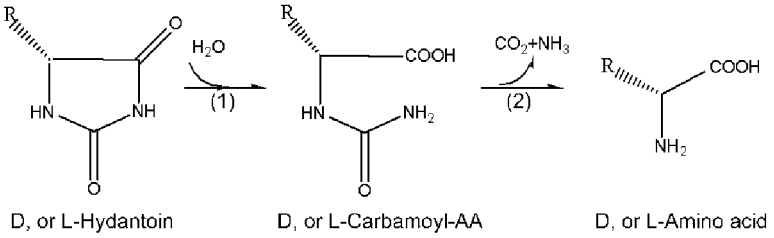


Figure 4. Production of optically pure amino acids by the combined action of hydantoinase and N-carbamoylase. (1) Hydrolysis of D,L-5-monosubstituted hydantoin by hydantoinase leads to the production of the corresponding carbamoyl amino acids (carbamoyl-AA). (2) With the liberation of CO_2, deamination of carbamoyl-AA by steroselective N-carbamoylase leads to the production of optically pure amino acids

5. FUTURE PERSPECTIVES

The relatively untapped potential of unusual amino acids in the pharmaceutical industry will provide the incentive for the development of HYD-based production processes. Rapid advances in the development of this production process can be expected given the continuing efforts in expanding the substrate spectrum of HYD, altering its enantioselectivity, and the improvement of its activity and thermostability. All of these goals can be achieved by the approach of directed evolution (May *et al*., 2000) and the rational design of protein structures based on the comparative analysis of molecular structures (Cheron *et al*., 2004). Meanwhile, the availability and efficiency of preparative methods for immobilized enzymes and immobilized whole cells should greatly increase the competitiveness and feasibility of this production process. The initial approach of the immobilization of recombinant *E. coli* with k-carrageenan for D-HPG production proved unsatisfactory (Chao *et al*., 1999). Owing to the low solubility of the substrate, the occurrence of substrate fouling on the gel surface made the immobilized cells inappropriate for reuse. Recently, ChBD-based cell immobilization has been explored by the construction of a tripartite gene fusion consisting of *lpp-ompA*-ChBD in *E. coli* (Wang and Chao, 2005). The result showed that ChBD-displaying cells exhibited highly specific and stable binding to chitin within a wide range of pH (5–8) and temperatures (15–37 °C). It will be very interesting to investigate the potential of this approach by fixing the whole cell on a chitin surface for D-HPG production.

REFERENCES

Chao, Y.P., Chiang, C.J. and Chen, P.T. (2000). Optimum ratio of D-carbamoylase to D-hydantionase for maximizing D-p-hydroxyphenylglycine productivity. Biotechnol. Lett. *22*, 99–103.

Chao, Y.P., Chiang, C.J., Lo, T.E. and Fu, H. (2000). Overproduction of D-hydantoinase and carbamoylase in a soluble form in *Escherichia coli*. Appl. Microbiol. Biotechnol. *54*, 348–353.

Chao, Y.P., Fu, H., Lo, T.E., Chen, P.T. and Wang, J.J. (1999). One-step production of D-p- hydroxyphenylglycine by recombinant *Escherichia coli* strains. Biotechnol. Prog. *15*, 1039–1045.

Cheon, Y.H., Kim, H.S., Han, K.H., Abendroth, J., Niefind, K., Schomburg, D., Wang, J. and Kim, Y. (2002). Crystal structure of D-hydnatoinase from *Bacillus stearothermophilus*: insight into the stereochemistry of enantioselectivity. Biochem. *41*, 9410–9417.

Cheron, Y.H., Park, H.S., Kim, H.S., Kim, Y. and Kim, H.S. (2004). Manipulation of the active site loops of D-hydnatoinase, a (β/α) 8-barrel protein, for modulation of the substrate specificity. Biochem. *43*, 7413–7420.

Chern, J.T. and Chao, Y.P. (2005). Chitin-binding domain based immobilization of D-hydantoinase. J. Biotechnol. *117*, 267–275.

Chiang, C.J, Chen, H.C., Kuo, H.F., Chao, Y.P. and Tzen, J.T. C. (2006). A simple and effective method to prepare immobilized enzymes using reconstituted oil bodies. Enzyme Microbial. Tech. *39*, 1152–1158.

Eadie, G.S., Bernheim, F. and Bernheim, M.L.C. (1949). The partial purification and properties of animal and plant hydantionase. J. Biol. Chem. *181*, 449–458.

Frandsen, G.I., Mundy, J. and Tzen, J.T.C. (2001). Oil bodies and their associated proteins, oleosin and caleosin. Physiol. Plant. *112*, 301–307.

Gaebler, O.H. and Keltch, A.K. (1926). On the metabolism of hydantoins and hydantoic acid. J. Biol. Chem. *70*, 763–777.

Hashimoto, M., Ikegami, T., Seino, S., Ohuchi, N., Fukada, H., Sugiyama, J., Shirakawa, M. and Watanabe, T. (2000). Expression and characterization of the chitin-binding domain of chintinase A1 from *Bacillus circulans* WL-12. J. Bacteriol. *182*, 3045–3054.

Huang, C.Y., Chao, Y.P. and Yang, Y.S. (2003). Purification of industrial hydantoinase in one chromatographic step without any tag. Protein Expr. Purif. *30*, 134–139.

Kim, G.J. and Kim, H.S. (1998). Identification of the structural similarity in the functionally related amidohydrolase acting on the cyclic amide ring. Biochem. J. *330*, 295–302.

Kim, G.J. Lee, D.E. and Kim, H.S. (2000). Construction and evaluation of a novel bifunctional N-carbamylase-D-hydantoinase fusion enzyme. Appl. Environ. Microbiol. *66*, 2133–2138.

May, O., Nguyen, P.T. and Arnold, F.H. (2000). Inverting enantioselectivity by directed evolution of hydnatoinase for improved production of L-methionine. Nat. Biotechnol. *18*, 317–320.

Olivieri, R. Fascetti, E., Angelini, L. and Degen, L. (1979). Enzymatic conversion of N-carbamoyl-D-amino acids to D-amino acids. Enzyme. Microb. Technol. *1*, 201–204.

Olivieri, R., Fascetti, E., Angellini, L. and Degen, L. (1981). Microbial transformation of racemic hydantoins to D-amino acids, Biotechnol. Bioeng. *23*, 2173–2183.

Owgawa, J. and Shimizu, S. (2002). Industrial microbial enzymes: their discovery by screening and use in large-scale production of useful chemicals in Japan. Curr. Opin. Biotechnol. *13*, 367–75.

Pietzsch, M., Wiese, A. Ragnitz, K, Wilms, B., Altenbuchner, J., Mattes, R. and Syldatk, C. (2000). Purification of recombinant hydantoinase and L-N-carbamoylase from *Arthrobacter surescens* expressed in *Escherichia coli*: comparison of wild-type and genetically modified proteins. J. Chromatogr. B Biomed. Sci. Appl. *737*, 179–186.

Thoden, J.B., Phillips, G.N. jr., Neal, T.M., Raushel, F.M. and Holden, H.M. (2001). Molecular structure of dihydroorotase: a paradigm for catalysis through the use of a binuclear metal center. Biochem. *40*, 6989–6997.

Vogels, G.D. and Van der Drift, C. (1976). Degradation of purines and pyrimidines by micro-organisms. Bacteriol. Rev. *40*, 403–68.

Wang, J.Y. and Chao, Y.P. (2006). Immobilization of cells with surface-displayed chitin-binding domain. Appl. Environ. Microbiol. *72*, 927–931.

Wilms, B., Wiese, A., Syldatk, C., Mattes, R. and Altenbuchner, J. (2001). Development of an *Escherichia coli* whole cell biocatalyst for the production of L-amino acids. J. Biotechnol. *86*, 19–30.

Xu, Z., Liu, Y., Yang, Y., Jiang, W., Arnold, E. and Ding, J. (2003). Crystal structure of D-hydantoinase from *Burkholderia pickettii* at a resolution of 2.7 angstroms: insight into the molecular basis of enzyme thermostability. J. Bacteriol. *185*, 4038–4049.

Yamada, Y., Takahashi, S., Kii, Y. and Kumagai, H. (1978). Distribution of hydantoin hydrolyzing activity in micro-organisms. J. Ferment. Technol. *56*, 484–491.

Yokozeki, K., Sano, K., Eguchi, C., Yamada, K. and Mitsugi, K. (1987). Enzymatic production of L-tryptophan from D,L-5-indolylmethylhydantoin by mutants of *Flavobacterium* species T-523. Agri. Biol. Chem. *51*, 819–825.

SUBJECT INDEX

ORGANISM INDEX