M. Zouhair Atassi *Editor*

Protein Reviews

Volume 17

Springer

Advances in Experimental Medicine and Biology

Protein Reviews

Volume 925

Subseries Editor
M. Zouhair Atassi

More information about this series at

M. Zouhair Atassi
Editor

Protein Reviews

Volume 17

Springer

Editor
M. Zouhair Atassi
Biochem and Mol Biol
Baylor College of Medicine
Houston, TX, USA

Preface

Protein Reviews is a book series which has been published by Springer since 2005. It has published 15 printed volumes to date (to see previous volumes, please go to http://www.springer.com/series/6876). Each volume was dedicated to a particular theme:

– *Viral Membrane Proteins: Structure, Function, and Drug Design*
– *The p53 Tumor Suppressor Pathway and Cancer*
– *Proteomics and Protein-Protein Interactions*
– *Protein Misfolding, Aggregation and Conformational Diseases: Part A*
– *Protein Interactions*
– *Protein Misfolding, Aggregation and Conformational Diseases: Part B*
– *Cell Stress Proteins*
– *Actin-Binding Proteins and Disease*
– *Dioxygen Binding and Sensing Proteins*
– *Ribonuclease P*
– *The Circadian Clock*
– *Post-Translational Modifications in Health and Disease*
– *Folding of Disulphide Proteins*
– *The Nucleolus*
– *cp*

From here on, to speed up the publication process and enhance accessibility, all articles will appear online before they are published in a printed book. The book series will appear as a subseries of Advances in Experimental Medicine and Biology (http://www.springer.com/series/14330) in volumes, each of which will focus on a given theme or volumes that contain reviews on an assortment of topics, in order to stay up to date and to publish timely reviews in an efficient manner.

The aim of *Protein Reviews* is to serve as a publication vehicle for reviews that focus on crucial contemporary and vital aspects of protein structure, function, evolution and genetics. Publications will be selected based on their importance to the understanding of biological systems, their relevance to the unravelling of issues associated with health and disease or their impact on scientific or technological advances and developments. Proteins linked to diseases or to the appearance and progress of diseases will obviously provide

essential topics that may be covered in this series. Moreover, proteins that are, or can be, used as potential biomarkers or as candidates for treatment and/or for the design of distinctive, new therapeutics will receive high attention in this book series.

The issues may include biochemistry, biophysics, immunology, structural and molecular biology, genetics, molecular and cellular mechanisms of action, clinical studies and new pioneering therapies. A given volume may be focused on a particular theme or may contain a selected assortment of different current topics.

The authors of the articles are selected from leading basic or medical scientists in academic or industrial organizations. The invited authors are nominated by the editorial board or by experts in the scientific community. However, interested individuals may suggest a topic for review and/or may propose a person to review a current important topic. Colleagues interested in writing a review or in guest-editing a special thematic issue are encouraged to submit their proposals and the list of authors of the suggested chapter/topics to the editor before submitting a manuscript.

The manuscripts are reviewed and evaluated in the usual manner by experts in the field. The articles will be published online no later than 6 weeks after editorial review and acceptance.

It should be re-emphasized that *Protein Reviews* will publish all accepted review articles online before they appear in print. And there will be no page or colour charges and no page or colour image limitations.

I hope that *Protein Reviews* will continue to serve the scientific community as a valuable vehicle for the dissemination of vital and essential contemporary discoveries on protein molecules and their immensely versatile biological activities.

M. Zouhair Atassi
Editor-in-Chief

Contents

Aggregation of FET Proteins as a Pathological Change in Amyotrophic Lateral Sclerosis

Yoshiaki Furukawa and Eiichi Tokuda

Abstract

Amyotrophic lateral sclerosis (ALS) is a fatal motor neuron disease that is characterized by the formation of abnormal inclusions in neurons. While the pathomechanism of ALS remains obscure, a number of proteins have been identified in the inclusion bodies, and the pathological roles of RNA-binding proteins have been increasingly emphasized. Among those, the FET proteins (FUS, EWSR1, TAF15) were recently identified as RNA-binding proteins in pathological inclusions of ALS and other neurodegenerative diseases; moreover, mutations in the genes encoding the FET proteins were found to be associated with familial forms of ALS. FET proteins are normally localized in the nucleus, but the introduction of pathogenic mutations in FET proteins leads to their abnormal redistribution to the cytoplasm, where they form aggregates. While further investigation will be required to understand the intracellular factors controlling the aggregation propensities of FET proteins, they are thought to lose their physiological functions and become toxic through their misfolding/aggregation. Here, we will briefly review recent advances of our understanding of the physiological functions and aggregation behavior of FET proteins *in vivo* as well as *in vitro*.

Keywords

FUS • FET family proteins • Amyotrophic lateral sclerosis • Protein aggregation • Neurodegenerative diseases

Y. Furukawa (✉) and E. Tokuda
Laboratory for Mechanistic Chemistry of Biomolecules,
Department of Chemistry, Keio University, 3-14-1
Hiyoshi, Kohoku, Yokohama 223-8522, Japan
e-mail: furukawa@chem.keio.ac.jp

1 Introduction

Amyotrophic lateral sclerosis (ALS) is a late-onset progressive disease with degeneration of motor neurons. Most of the cases (>90 %) are sporadic, and the pathomechanism remains

obscure. No cures are currently available for this devastating disease, but increasing numbers of genes responsible for familial forms of ALS (fALS) have been recently identified (http://alsod.iop.kcl.ac.uk). Mutations in *SOD1* gene were first identified as a cause of fALS in 1993 (Rosen et al. 1993), and transgenic rodents expressing SOD1 with pathogenic mutations have long served as a standard model in the research of ALS (Turner and Talbot 2008). In 2006, Arai et al. (2006) and Neumann et al. (2006) reported milestone papers where a DNA/RNA-binding protein, TDP-43, was identified as a constituent of pathological inclusions in sporadic ALS patients. Soon after the identification, several ALS-causing mutations were successively reported in the *TARDBP* gene, which encodes for the TDP-43 protein (Gitcho et al. 2008; Kabashi et al. 2008; Sreedharan et al. 2008; Van Deerlin et al. 2008). While mutations in *TARDBP* have been described in ~4 % of familial and ~1.5 % of sporadic ALS cases, the neuropathology in most cases of sporadic ALS is characterized by the abnormal cytoplasmic accumulation of TDP-43 in neurons and glia (Mackenzie et al. 2010a). Importantly, TDP-43-positive cytoplasmic inclusions have been characterized in almost half of the frontotemporal lobar degeneration (FTLD) cases (Mackenzie et al. 2010a). Thus, TDP-43 has become the most important player as a pathological protein in ALS.

In 2009, mutations in a gene encoding for Fused in Sarcoma (FUS) were found to be causative in a subset of ALS cases (ALS-*FUS*) (Kwiatkowski et al. 2009; Vance et al. 2009). Given that FUS is also a DNA/RNA-binding protein, many researchers have now focused upon the possible involvement of DNA/RNA metabolism in the pathomechanism of ALS. ALS-*FUS* accounts for ~4 % of familial and <1 % of sporadic ALS cases, and its neuropathology has been characterized by abnormal FUS-positive inclusions in the cytoplasm of neurons and glia (Mackenzie et al. 2010a). FUS has also been identified as a component of the inclusions in atypical frontotemporal lobar degeneration with ubiquitin-positive inclusions

(FTLD-U), neuronal intermediate filament inclusion disease (NIFID), and basophilic inclusions body disease (BIBD), which are collectively called FTLD-FUS (Mackenzie et al. 2010b). No genetic abnormalities of *FUS* have been identified in FTLD-FUS, and it remains unknown how FUS facilitates these pathological changes.

As described below, both TDP-43 and FUS have a "prion-like" domain with low amino acid complexity (Gitler and Shorter 2011; Polymenidou and Cleveland 2011). Prion proteins in mammals (PrP) and yeast (*e.g.* Sup35) are equipped with a low complexity region, which is essential for the formation of fibrillar aggregates and also for the infectivity and transmissibility as prions (Chien et al. 2004). The high aggregation propensities of the prion-like domains in TDP-43 and FUS are consistent with their involvement in pathological inclusions observed in motorneurons of ALS patients, which has spurred many researchers to investigate the aggregation mechanism of TDP-43/FUS. Very interestingly, the other protein candidates to cause ALS were surveyed by systematic screening of RNA-binding proteins using yeast, among which missense mutations in *TAF15, EWSR1* and *hnRNPA1* were indeed identified in ALS patients (Couthouis et al. 2011, 2012; Kim et al. 2013). All of these proteins have a prion-like domain with low complexity and have a domain organization which is very similar to that of TDP-43 (hnRNPA1) or FUS (TAF15, EWSR1). So, there may be a common pathway for those RNA-binding proteins to gain some toxicity that leads to ALS. In this review, we will focus upon the FET proteins (FUS, EWSR1, and TAF15) and briefly summarize recent advance of our understanding of their pathological roles in ALS.

2 Structure and Physiological Functions of FET Proteins

FUS, EWSR1, and TAF15 constitute the FET protein family and are normally localized in the nucleus (Tan and Manley 2009). While it remains obscure if there is specificity or

redundancy in the functions of different FET proteins, all three FET proteins are commonly involved in several types of sarcomas such as Ewing sarcomas and myxoid liposarcomas, where genetic translocation of FET proteins form fusion oncoproteins (Riggi et al. 2007). Also, the FET proteins have the same domain organization: an N-terminal region with low amino acid complexity that is rich in Gln, Gly, Ser, and Tyr (LC domain), an RNA recognition motif (RRM), a Zn-finger motif (ZnF), three regions rich in Arg and Gly (RGG1, 2, and 3), and a nuclear localization signal called PY-NLS at the C-terminus (Schwartz et al. 2015). Many ALS-causing mutations are reported in the PY-NLS of FUS (Mackenzie et al. 2010a), which retards the nuclear transporting function and thus results in the abnormal accumulation of mutant FUS in the cytoplasm. Very interestingly, the affinity of PY-NLS with its nuclear import receptor, transportin, has been reported to significantly decrease with a series of ALS-causing mutations and appears to inversely correlate with the disease severity (Dormann et al. 2012; Zhang and Chook 2012), while the cytoplasmic accumulation of FUS with mutations in the regions outside PY-NLS has not been well described. Instead, mutations in RGG1 of FUS, RGG2/3 of EWSR1, and RGG3 of TAF15 have been proposed to accelerate the protein aggregation (Couthouis et al. 2011, 2012; Nomura et al. 2013); however, because FET proteins are quite prone to aggregation, effects of mutations on aggregation kinetics need to be investigated in more detail.

FET proteins can bind single-stranded DNA (ssDNA), double-stranded DNA (dsDNA), and RNA. FUS has been reported to bind ssDNA more tightly than dsDNA *in vitro* (Schwartz et al. 2015) and shows the ability to anneal DNA strands and promote ssDNA invasion of dsDNA (Bertrand et al. 1999; Baechtold et al. 1999). The apparent annealing activity of FET proteins may thus be involved in the DNA break repair; indeed, a knockout of the *FUS* gene in mice has been shown to accumulate DNA breaks after treatment with ionizing radiation (Kuroda et al. 2000). Furthermore, FUS is recruited to DNA lesions by laser irradiation (Wang et al. 2013), and phosphorylated FUS can bind dsDNA breaks and Holliday junctions (Gardiner et al. 2008). Nonetheless, it largely remains obscure whether the FET proteins directly interact with DNA *in vivo*, and physiological roles of the DNA binding by FET proteins need to be investigated in more detail.

In contrast to the DNA binding, the interaction of FET proteins with RNA is more established *in vivo* as well as *in vitro*. As described above, RRM, ZnF, and RGG motifs in FET proteins play major roles in the binding of RNA (Schwartz et al. 2015). While the RRM alone exhibited very weak affinity to RNA, inclusion of the flanking RGG motifs significantly strengthened the RNA binding of RRM (Liu et al. 2013; Schwartz et al. 2013; Lerga et al. 2001). FET proteins appear to exhibit no strong sequence-specificity of the substrate RNAs (Schwartz et al. 2015), but the binding sites of FUS in mRNA are enriched in 5′ untranslated regions (UTRs) (Ishigaki et al. 2012; Lagier-Tourenne et al. 2012). Also, FUS has been proposed to preferentially bind introns in pre-mRNA with a saw tooth pattern, in which the binding of FUS occurs most at a 5′ splice site but least at a 3′end of the intron (Rogelj et al. 2012). Actually, preferential binding of FUS at 5′-UTRs and introns is consistent with its major physiological functions in transcription and RNA processing.

FET proteins have been proposed to directly bind transcription factors and splicing factors/pre-mRNA to control the transcription and the RNA processing, respectively (Schwartz et al. 2015). Alternatively, FET proteins have been shown to oligomerize and form fibrils through their LC domains, and the fibrils of FET proteins can bind RNA polymerase II (RNA Pol II) and then trigger transcription (Schwartz et al. 2013; Kwon et al. 2013). Splicing and polyadenylation of RNA are considered to occur co-transcriptionally; therefore, fibrils of FET proteins would provide a molecular framework that can facilitate the interactions among RNA Pol II, splicing factors, and pre-mRNA (Schwartz et al. 2015). Because oligomerization of FET proteins is suppressed by phosphorylation

at their LC domains (Han et al. 2012), the levels of mRNA transcripts could be controlled by the post-translational modification of FET proteins. In some types of sarcomas, furthermore, genetic translocation of FET proteins is known to produce oncoproteins in which their LC domains are fused to a DNA-binding domain from another transcription factor such as ERG, CHOP or FLI1 (Tan and Manley 2009; Riggi et al. 2007). Such abnormally fused proteins can act as transcriptional activators and are involved in tumorigenesis.

FET proteins are predominantly localized in the nucleus so as to regulate the transcription and the RNA processing, but a small fraction of the FET proteins is exported from the nucleus and shuttled to the cytoplasm (Zinszner et al. 1997). The nuclear-cytoplasmic shuttling of FET proteins would occur in association with mRNA, and once exported to the cytoplasm, the FET proteins dissociate from its bound mRNA and then return to the nucleus. As described above, the FET proteins possess a PY-NLS at the C-terminus that facilitates the nuclear import through an interaction with transportin (Lee et al. 2006). While the regulation of the nuclear-cytoplasmic shuttling of the FET proteins remains obscure, the nuclear import of the FET proteins is negatively regulated by chemical modifications near the PY-NLS, such as the methylation of Arg (Dormann et al. 2012; Belyanskaya et al. 2003) and the phosphorylation of Tyr (Leemann-Zakaryan et al. 2011). Given that cytoplasmic accumulation of FUS is observed in the ALS-*FUS* cases, post-translational modifications affecting the function of PY-NLS may also have pathological roles in ALS (Dormann et al. 2012; Suarez-Calvet et al. 2016). Also, notably, significant concentrations of FUS are observed in the dendritic spines of neurons, where mRNAs are stored for their local translation upon synaptic activity (Fujii et al. 2005). Reduction of dendritic spines has been reported in FUS-knockout mice; therefore, FUS plays important roles in the trafficking of mRNA along dendrites, which would be affected by ALS-causing mutations.

In the cytoplasm, FET proteins are involved in the formation of stress granules containing non-translating mRNAs, translation initiation components, and RNA-binding proteins (Andersson et al. 2008). In response to a variety of environmental stresses, the translation of mRNAs that are not essential for survival is stalled by specific recruitment of these mRNAs into the stress granules (Anderson and Kedersha 2009). After the removal of stresses, the stress granules dissolve, and the stalled translation of mRNAs in stress granules restarts. Notably, participation of FUS in the formation of cytoplasmic stress granules has been confirmed only when osmotic stress but not when other types of stresses such as oxidative stress were applied (Sama et al. 2013). In contrast, mutations in the PY-NLS increases the cytoplasmic fraction of FUS and leads to the accumulation of FUS in stress granules in response to multiple stresses (Bosco et al. 2010). Such accumulation of FUS in stress granules hence increases its local concentration and may trigger the formation of fibrillar aggregates. Also, formation of insoluble FUS aggregates results in the persistence of stress granules even after the stress is removed (Bosco et al. 2010; Dormann et al. 2010). Neuronal FUS-positive cytoplasmic inclusions in a fALS case with a mutation in *FUS* (R521C) and also in sporadic FTLD-FUS cases are clearly co-localized with the stress granule markers, PABP-1 and eIF4G (Dormann et al. 2010). Neuronal cytoplasmic inclusions in the FTLD cases with TDP-43 pathologies were not stained with those stress granule markers (Dormann et al. 2010); therefore, stress granules may have roles in the formation of inclusion bodies containing FUS proteins.

3 Pathological Inclusions Containing FET Proteins in ALS Cases

As mentioned above, ALS-*FUS* cases have been characterized by the FUS-positive cytoplasmic inclusions in neurons and glia, but their morphologies appear to be heterogeneous and dependent upon mutations (Mackenzie et al. 2011); for example, round and tangle-like inclusions have distinctly characterized the ALS

cases with P525L and R521C mutations in *FUS* gene, respectively. While the P525L mutation exhibits significantly earlier onset of disease than the R521C mutation, more investigation will be required to test if morphologies of inclusions correlates with disease phenotypes. Absence of FUS-positive inclusions was reported in sporadic ALS cases, but it has also been reported that FUS-immunoreactivity is affected by antigen retrieval methods (Deng et al. 2010). When the antigen retrieval was performed by the high-pressure decloaking chamber, FUS-immunoreactivity was found to become evident in the inclusions formed in sporadic and familial ALS cases, except for those with *SOD1* mutations (Deng et al. 2010; Keller et al. 2012). Like TDP-43, therefore, FUS may be a common component of cytoplasmic inclusions in non-SOD1 ALS cases.

In ALS-*FUS* cases, cytoplasmic inclusions have been reported to show immunoreactivity to stress granule markers (PABP-1 and eIF4G) (Dormann et al. 2010), Rho guanine nucleotide exchange factor (RGNEF), TDP-43 (Keller et al. 2012), p62 (Dormann et al. 2010; Keller et al. 2012), ubiquitin (Keller et al. 2012), peripherin (Keller et al. 2012), optineurin (Keller et al. 2012; Ito et al. 2011), and ataxin-2 (Farg et al. 2013). Among those, it is notable that ALS-causing/ALS-risk mutations are identified in genes encoding TDP-43 (Gitcho et al. 2008; Kabashi et al. 2008; Sreedharan et al. 2008; Van Deerlin et al. 2008), peripherin (Gros-Louis et al. 2004), optineurin (Maruyama et al. 2010), and ataxin-2 (Elden et al. 2010). Furthermore, even though FUS did not co-localize with SOD1 in cytoplasmic inclusions (Deng et al. 2010), the presence of misfolded SOD1 has been suggested in an ALS-*FUS* case through an unknown mechanism (Pokrishevsky et al. 2012). While increasing numbers of genes have been recently identified as a cause of fALS (Robberecht and Philips 2013), there might be a unified molecular pathomechanism among fALS cases with mutations in distinct genes.

No immunohistochemical examination has been performed on ALS cases with mutations in the other FET proteins, EWSR1 and TAF15. In sporadic ALS cases, EWSR1 and TAF15 were shown to be present in cytoplasmic puncta of spinal cord neurons and/or in a diffuse pattern throughout the cytoplasm (Couthouis et al. 2011, 2012). While co-localization of EWSR1/TAF15 with FUS has not been fully examined, inclusions immunoreactive to EWSR1 and TAF15 were reported in an ALS-FUS case (Takeuchi et al. 2013). In contrast, co-localization of EWSR1/TAF15 with FUS-positive inclusions was not confirmed in ALS-*FUS* cases but was evident in FTLD-FUS cases, including atypical FTLD-U, NIFID and BIBD (Neumann et al. 2011). Such distinct involvement of EWSR1 and TAF15 in the FUS-positive inclusions is surprising, because those three proteins, FUS, EWSR1, and TAF15, have very similar structures with high aggregation propensities. Also, transportin, which is a nuclear import receptor for FET proteins, accumulates in FUS-positive inclusions in FTLD-FUS (Neumann et al. 2012) and ALS without *FUS* mutations (Takeuchi et al. 2013) but is absent in the inclusions in ALS-*FUS* cases (Neumann et al. 2012). It has also been shown that the Arg methylation plays distinct pathological roles in FTLD-FUS and ALS-*FUS*. In ALS-*FUS* but not FTLD-FUS cases, the pathological inclusions contained FUS with asymmetrically dimethylated arginines (ADMA) (Dormann et al. 2012). In contrast, unmethylated/monomethylated Arg (UMA/MMA) at RGG3 region of FUS constituted the inclusions observed in FTLD-FUS but not in ALS-*FUS* cases (Suarez-Calvet et al. 2016). Methylations of Arg residues in FUS have been shown to reduce its binding to transportin and thus suppress the nuclear import of FUS (Dormann et al. 2012; Suarez-Calvet et al. 2016). This is consistent with the observation described above that transportin is recruited into the FUS-positive inclusions in FTLD-FUS but not ALS-*FUS* cases. Alterations in the regulation of Arg methylation of FUS would explain the distinct pathogenesis of FTLD and ALS.

ALS-*FUS* is characterized by the FUS-positive inclusions in neurons and glia, but any biochemical changes in mutant FUS isolated from human tissues have not been examined.

Western blotting analysis on the spinal cord homogenates from only one ALS-*FUS* case with the R521L mutation in *FUS* has been reported, and no modifications/truncations in FUS were confirmed with electrophoresis (Deng et al. 2010). To our knowledge, there are currently no reports showing that pathogenic mutations decrease the solubility of FUS proteins in human tissues. In contrast, several papers have supported decreased solubility of FUS in FTLD-FUS cases, including atypical FTLD-U, BIBD, and NIFID. Compared to non-neurological controls, higher levels of FUS in FTLD-FUS cases were insoluble in RIPA buffer (50 mM Tris, 150 mM NaCl, 5 mM EDTA, 1 % NP-40, 0.5 % sodium deoxycholate, and 0.1 % SDS) but soluble in either 2 % SDS or 7 M urea (Neumann et al. 2009, 2011, 2012; Lashley et al. 2011; Page et al. 2011). While insolubilization of FUS proteins in FTLD-FUS cases was statistically significant, the level of insoluble FUS appeared to correlate with the severity of FUS pathologies and were even comparable to those of controls in some cases (Neumann et al. 2009). Moreover, FUS-positive inclusions in FTLD-FUS were thioflavin-S negative (Lashley et al. 2011; Bigio et al. 2013), suggesting distinct structural/biochemical properties from those of insoluble amyloids rich in β-sheet structures. More biochemical characterization on pathological changes of FUS will thus be required in FTLD-FUS and ALS-*FUS* cases, in particular.

FUS has been also reported to be a component of inclusions formed by proteins with a polyglutamine (polyQ) stretch. Abnormal expansion of a polyQ stretch in Huntingtin, Ataxin 1, Ataxin 2, Ataxin 3, and Atrophin 1 leads to the formation of neuronal intranuclear inclusions in Huntington's disease (HD), spinocerebellar ataxia (SCA) 1, SCA 2, SCA 3, and dentatorubral-pallidoluysian atrophy (DRPLA), respectively (Zoghbi and Orr 2000). Actually, before the identification of *FUS* as a cause of fALS, FUS was first characterized as a protein recruited into the neuronal intranuclear inclusions in HD (Doi et al. 2008); furthermore, EWSR1 and TAF15 were also identified as components of insoluble aggregates formed in a HD model mouse expressing Huntingtin with an abnormally expanded polyQ (Doi et al. 2008). Later, recruitment of FUS into neuronal intranuclear inclusions has been confirmed also in the other polyQ diseases, SCA 1, 2, 3, and DRPLA (Doi et al. 2010; Woulfe et al. 2010). In contrast, FUS pathologies were not observed in FTLD with TDP-43 pathologies, multiple system atrophy, fragile X-associated tremor/ataxia syndrome, oculopharyngeal muscular dystrophy, and dementia with Lewy bodies (Doi et al. 2010; Woulfe et al. 2010). Recruitment of FUS proteins into intramolecular inclusions is thus considered to occur specifically in polyQ diseases. Despite this, the pathological relevance of FET proteins in those polyQ diseases remains to be determined.

4 *In vitro* Aggregation of Recombinant FET Proteins

FET proteins are quite highly prone to aggregation and difficult to handle while expressed in *E. coli* and purified *in vitro*. In many of aggregation-prone proteins, fusion with glutathione-S-transferase (GST) can increase their solubility; indeed, soluble FUS proteins can be prepared when tagged with an N-terminal but not C-terminal GST (Sun et al. 2011). As shown by our group, however, most fractions of GST-FUS expressed in *E. coli* were collected as insoluble inclusions, and significant amounts of the truncated GST-FUS were also contaminated even after the affinity purification using glutathione Sepharose resins (Nomura et al. 2013). We thus further attached a 6x His tag at the C-terminus of GST-FUS and then successfully removed the truncated proteins by additional purification using Ni^{2+}-affinity resins. Furthermore, purified GST-FUS-His samples were found to contain *E. coli* endogenous nucleotides (Nomura et al. 2013); therefore, dissociation of those nucleotides with a buffer containing high concentrations of salts (1 M NaCl) was required to prepare nucleotide-free GST-FUS-His samples.

As described above, FET proteins are characterized by the N-terminal LC domain rich in Q, S, Y, and G. Such biased compositions in amino acid residues have been shown to increase the aggregation propensity of proteins; an extreme example for this is a protein with a polyglutamine tract (Scherzinger et al. 1997). Indeed, Sun et al. have reported that FUS (a.a. 1 – 526) becomes resistant to aggregation by deleting its N-terminal LC domain with a part of the RGG1 domain (a.a. 1 – 266) (Sun et al. 2011). Quite curiously, however, the deleted LC region with an adjacent C-terminal sequence (a.a. 1 – 373) also remained soluble (Sun et al. 2011). Furthermore, the C-terminally truncated FUS (a.a. 1 – 422) was detected in the insoluble pellets, which were difficult to detect by turbidity (Sun et al. 2011), while aggregates of both FUS and the truncated FUS (a.a. 1 – 422) exhibited amyloid-like fibrillar morphologies. In the paper reported by the other group (Kato et al. 2012), nonetheless, the LC domain of FUS (a.a. 1 – 214) fused with a fluorescent protein was clearly shown to form amyloid-like fibrils with cross-β structure, which further constituted hydrogels. It is also notable that distinct LC domains from different RNA-binding proteins (RBM3, hnRNPA1/A2, CPEB2, and TIA1) form hydrogels (Kato et al. 2012); therefore, the LC domain is considered to describe high aggregation propensities of FUS. Despite this, the fibrillar aggregates of FUS did not show reactivity to Thioflavin T and were solubilized with a buffer containing a detergent, SDS, both of which are distinct properties from those of *bona fide* amyloid fibrils (Sun et al. 2011; Kato et al. 2012).

In contrast, our group has shown that a full-length FUS protein dual-tagged with an N-terminal GST and a C-terminal 6x His remained soluble and we also found that an ALS-causing mutation, G156E, triggered the formation of insoluble aggregates with amyloid-like fibrillar morphologies (Nomura et al. 2013). Furthermore, the fibrils of GST-FUS(G156E)-His reacted with Thioflavin T, suggesting distinct properties from those of the hydrogels prepared from the LC domain as mentioned above. While it remains to be established how Thioflavin T

becomes fluorescent upon the binding with amyloid fibrils, the distinct reactivity with Thioflavin T would be somehow caused by structural differences between hydrogels and our insoluble aggregates and hence reflect structural heterogeneity/diversity of FUS fibrils. ALS-causing mutations at the C-terminal domain of FUS (H517Q, R521H, R521C, P525L) were, however, found to have little effects on the aggregation reaction of FUS, and also, GST-FUS-His with pathogenic mutations in the RGG1 region (G225V and M254V) did not form insoluble aggregates (Nomura et al. 2013; Sun et al. 2011). Given that FUS formed insoluble aggregates by artificial mutations of G144E, G154E, and G156D, introduction of acidic residues in the LC domain would facilitate the formation of aggregation-prone conformations of FUS (Nomura et al. 2013). Based upon those *in vitro* studies, we suspect that ALS patients with the G156E mutation in *FUS* might be characterized by pathologies distinct from those with the other *FUS* mutations. Indeed, expression of FUS with the G156E mutation in primary cultured neurons resulted in the formation of intranuclear foci, supporting the increased propensity of FUS aggregation by the G156E mutation (Nomura et al. 2013). Very recently, furthermore, the LC domain of FUS (a.a. 2 – 214) with an N-terminal His tag has been examined for the formation of hydrogels *in vitro*; the LC domain becomes condensed into gels at 4 °C but is re-dissolved into a clear solution upon re-warming to 23 °C (Murakami et al. 2015). After several cycles of varying temperature between 4 and 23 °C, the LC domain becomes irreversibly gelled, and such irreversible gels exhibit fibrillar morphologies with poor solubility in RIPA buffer. Less numbers of temperature cycles were required to make irreversible gels of the LC domain with the G156E mutation, supporting our proposal that the G156E mutation increases the aggregation propensities of FUS.

The other FET proteins, EWSR1 and TAF15, have also been examined for their aggregation propensities *in vitro* (Couthouis et al. 2011, 2012). Proteolytic cleavage of an N-terminal GST tag from GST-EWSR1 and GST-TAF15

proteins commenced the increase of solution turbidity, and the resultant aggregates exhibited amyloid-like fibrillar morphologies. Aggregation of FET proteins monitored by solution turbidity was then found to be accelerated by pathogenic mutations, G511A and P552L in the RGG3 region of EWSR1, and G391E and R408C in the RGG3 region of TAF15 (Couthouis et al. 2011, 2012). In addition to the C-terminal PY-NLS, nuclear localization of EWSR1 is also controlled by the RGG3 region (Shaw et al. 2009), where the pathogenic mutations (G511A and P552L) increased accumulation of EWSR1 in the cytoplasm (Couthouis et al. 2012). Unlike FUS and EWSR1, localization of exogenously expressed TAF15 was not strictly restricted to the nucleus in rat primary embryonic neurons; rather, the presence of cytoplasmic TAF15 was confirmed even in the absence of mutations (Couthouis et al. 2011). Also, pathogenic mutations did not lead to an increased fraction of cytoplasmic TAF15, but mutant TAF15 in the cytoplasm was found to accumulate in foci, consistent with the *in vitro* observation that pathogenic mutations increased the aggregation propensities of TAF15 (Couthouis et al. 2011). It remains unknown how mutations in the RGG region trigger the aggregation in all three FET proteins, and further investigation will also be required to reveal the significance of cytoplasmic mislocalization of the FET proteins in the formation of pathological aggregates/inclusions.

5 Concluding Remarks

As described above, abnormalities of FET proteins with or without mutations have been observed as pathological changes in ALS, FTLD and several other neurodegenerative diseases. Recently, a study on conditional human FUS transgenic mice and conditional FUS knockout mice has supported a disease model in which mutant FUS causes degeneration of motor neurons through a gain of toxicities but not a loss of physiological functions (Sharma et al. 2016). Conformational abnormalities of mutant FET proteins leading to aggregation would be related to the gain of such toxicities;

however, it is still enigmatic how pathogenic mutations increase the aggregation propensities of wild-type FET proteins, and also, it remains unknown why highly aggregation-prone FET proteins remain soluble and diffused in the nucleus under normal conditions. In other words, it is important to reveal a mechanism of how the cytoplasmic translocation of nuclear FET proteins are linked to their formation of aggregates. Deciphering the secret to suppress the FET aggregation in the nucleus will help in preventing the pathological changes of FUS proteins and eventually in developing cures for ALS.

Conflicts of Interest The authors declare no conflict of interest.

References

Anderson P, Kedersha N (2009) RNA granules: post-transcriptional and epigenetic modulators of gene expression. Nat Rev Mol Cell Biol 10:430–436

Andersson MK, Stahlberg A, Arvidsson Y, Olofsson A, Semb H, Stenman G, Nilsson O, Aman P (2008) The multifunctional FUS, EWS and TAF15 proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading and stress response. BMC Cell Biol 9:37

Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y, Oda T (2006) TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem Biophys Res Commun 351:602–611

Baechtold H, Kuroda M, Sok J, Ron D, Lopez BS, Akhmedov AT (1999) Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein TLS/FUS and is able to promote D-loop formation. J Biol Chem 274:34337–34342

Belyanskaya LL, Delattre O, Gehring H (2003) Expression and subcellular localization of Ewing sarcoma (EWS) protein is affected by the methylation process. Exp Cell Res 288:374–381

Bertrand P, Akhmedov AT, Delacote F, Durrbach A, Lopez BS (1999) Human POMp75 is identified as the pro-oncoprotein TLS/FUS: both POMp75 and POMp100 DNA homologous pairing activities are

associated to cell proliferation. Oncogene 18:4515–4521

Bigio EH, Wu JY, Deng HX, Bit-Ivan EN, Mao Q, Ganti R, Peterson M, Siddique N, Geula C, Siddique T, Mesulam M (2013) Inclusions in frontotemporal lobar degeneration with TDP-43 proteinopathy (FTLD-TDP) and amyotrophic lateral sclerosis (ALS), but not FTLD with FUS proteinopathy (FTLD-FUS), have properties of amyloid. Acta Neuropathol 125:463–465

Bosco DA, Lemay N, Ko HK, Zhou H, Burke C, Kwiatkowski TJ Jr, Sapp P, McKenna-Yasek D, Brown RH Jr, Hayward LJ (2010) Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. Hum Mol Genet 19:4160–4175

Chien P, Weissman JS, DePace AH (2004) Emerging principles of conformation-based prion inheritance. Annu Rev Biochem 73:617–656

Couthouis J, Hart MP, Shorter J, DeJesus-Hernandez M, Erion R, Oristano R, Liu AX, Ramos D, Jethava N, Hosangadi D, Epstein J, Chiang A, Diaz Z, Nakaya T, Ibrahim F, Kim HJ, Solski JA, Williams KL, Mojsilovic-Petrovic J, Ingre C, Boylan K, Graff-Radford NR, Dickson DW, Clay-Falcone D, Elman L, McCluskey L, Greene R, Kalb RG, Lee VM, Trojanowski JQ, Ludolph A, Robberecht W, Andersen PM, Nicholson GA, Blair IP, King OD, Bonini NM, Van Deerlin V, Rademakers R, Mourelatos Z, Gitler AD (2011) A yeast functional screen predicts new candidate ALS disease genes. Proc Natl Acad Sci U S A 108:20881–20890

Couthouis J, Hart MP, Erion R, King OD, Diaz Z, Nakaya T, Ibrahim F, Kim HJ, Mojsilovic-Petrovic J, Panossian S, Kim CE, Frackelton EC, Solski JA, Williams KL, Clay-Falcone D, Elman L, McCluskey L, Greene R, Hakonarson H, Kalb RG, Lee VM, Trojanowski JQ, Nicholson GA, Blair IP, Bonini NM, Van Deerlin VM, Mourelatos Z, Shorter J, Gitler AD (2012) Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis. Hum Mol Genet 21:2899–2911

Deng HX, Zhai H, Bigio EH, Yan J, Fecto F, Ajroud K, Mishra M, Ajroud-Driss S, Heller S, Sufit R, Siddique N, Mugnaini E, Siddique T (2010) FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis. Ann Neurol 67:739–748

Doi H, Okamura K, Bauer PO, Furukawa Y, Shimizu H, Kurosawa M, Machida Y, Miyazaki H, Mitsui K, Kuroiwa Y, Nukina N (2008) RNA-binding protein TLS is a major nuclear aggregate-interacting protein in huntingtin exon 1 with expanded polyglutamine-expressing cells. J Biol Chem 283:6489–6500

Doi H, Koyano S, Suzuki Y, Nukina N, Kuroiwa Y (2010) The RNA-binding protein FUS/TLS is a common aggregate-interacting protein in polyglutamine diseases. Neurosci Res 66:131–133

Dormann D, Rodde R, Edbauer D, Bentmann E, Fischer I, Hruscha A, Than ME, Mackenzie IR, Capell A, Schmid B, Neumann M, Haass C (2010) ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. EMBO J 29:2841–2857

Dormann D, Madl T, Valori CF, Bentmann E, Tahirovic S, Abou-Ajram C, Kremmer E, Ansorge O, Mackenzie IR, Neumann M, Haass C (2012) Arginine methylation next to the PY-NLS modulates Transportin binding and nuclear import of FUS. EMBO J 31:4258–4275

Elden AC, Kim HJ, Hart MP, Chen-Plotkin AS, Johnson BS, Fang X, Armakola M, Geser F, Greene R, Lu MM, Padmanabhan A, Clay-Falcone D, McCluskey L, Elman L, Juhr D, Gruber PJ, Rub U, Auburger G, Trojanowski JQ, Lee VM, Van Deerlin VM, Bonini NM, Gitler AD (2010) Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature 466:1069–1075

Farg MA, Soo KY, Warraich ST, Sundaramoorthy V, Blair IP, Atkin JD (2013) Ataxin-2 interacts with FUS and intermediate-length polyglutamine expansions enhance FUS-related pathology in amyotrophic lateral sclerosis. Hum Mol Genet 22:717–728

Fujii R, Okabe S, Urushido T, Inoue K, Yoshimura A, Tachibana T, Nishikawa T, Hicks GG, Takumi T (2005) The RNA binding protein TLS is translocated to dendritic spines by mGluR5 activation and regulates spine morphology. Curr Biol 15:587–593

Gardiner M, Toth R, Vandermoere F, Morrice NA, Rouse J (2008) Identification and characterization of FUS/TLS as a new target of ATM. Biochem J 415:297–307

Gitcho MA, Baloh RH, Chakraverty S, Mayo K, Norton JB, Levitch D, Hatanpaa KJ, White CL 3rd, Bigio EH, Caselli R, Baker M, Al-Lozi MT, Morris JC, Pestronk A, Rademakers R, Goate AM, Cairns NJ (2008) TDP-43 A315T mutation in familial motor neuron disease. Ann Neurol 63:535–538

Gitler AD, Shorter J (2011) RNA-binding proteins with prion-like domains in ALS and FTLD-U. Prion 5:179–187

Gros-Louis F, Lariviere R, Gowing G, Laurent S, Camu W, Bouchard JP, Meininger V, Rouleau GA, Julien JP (2004) A frameshift deletion in peripherin gene associated with amyotrophic lateral sclerosis. J Biol Chem 279:45951–45956

Han TW, Kato M, Xie S, Wu LC, Mirzaei H, Pei J, Chen M, Xie Y, Allen J, Xiao G, McKnight SL (2012) Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell 149:768–779

Ishigaki S, Masuda A, Fujioka Y, Iguchi Y, Katsuno M, Shibata A, Urano F, Sobue G, Ohno K (2012) Position-dependent FUS-RNA interactions regulate alternative splicing events and transcriptions. Sci Rep 2:529

Ito H, Fujita K, Nakamura M, Wate R, Kaneko S, Sasaki S, Yamane K, Suzuki N, Aoki M, Shibata N, Togashi S, Kawata A, Mochizuki Y, Mizutani T, Maruyama H, Hirano A, Takahashi R, Kawakami H, Kusaka H (2011) Optineurin is co-localized with FUS in basophilic inclusions of ALS with FUS mutation and in basophilic inclusion body disease. Acta Neuropathol 121:555–557

Kabashi E, Valdmanis PN, Dion P, Spiegelman D, McConkey BJ, Vande Velde C, Bouchard JP, Lacomblez L, Pochigaeva K, Salachas F, Pradat PF, Camu W, Meininger V, Dupre N, Rouleau GA (2008) TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. Nat Genet 40:572–574

Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J, Grishin NV, Frantz DE, Schneider JW, Chen S, Li L, Sawaya MR, Eisenberg D, Tycko R, McKnight SL (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell 149:753–767

Keller BA, Volkening K, Droppelmann CA, Ang LC, Rademakers R, Strong MJ (2012) Co-aggregation of RNA binding proteins in ALS spinal motor neurons: evidence of a common pathogenic mechanism. Acta Neuropathol 124:733–747

Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, MacLea KS, Freibaum B, Li S, Molliex A, Kanagaraj AP, Carter R, Boylan KB, Wojtas AM, Rademakers R, Pinkus JL, Greenberg SA, Trojanowski JQ, Traynor BJ, Smith BN, Topp S, Gkazi AS, Miller J, Shaw CE, Kottlors M, Kirschner J, Pestronk A, Li YR, Ford AF, Gitler AD, Benatar M, King OD, Kimonis VE, Ross ED, Weihl CC, Shorter J, Taylor JP (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature 495:467–473

Kuroda M, Sok J, Webb L, Baechtold H, Urano F, Yin Y, Chung P, de Rooij DG, Akhmedov A, Ashley T, Ron D (2000) Male sterility and enhanced radiation sensitivity in TLS(-/-) mice. EMBO J 19:453–462

Kwiatkowski TJ Jr, Bosco DA, Leclerc AL, Tamrazian E, Vanderburg CR, Russ C, Davis A, Gilchrist J, Kasarskis EJ, Munsat T, Valdmanis P, Rouleau GA, Hosler BA, Cortelli P, de Jong PJ, Yoshinaga Y, Haines JL, Pericak-Vance MA, Yan J, Ticozzi N, Siddique T, McKenna-Yasek D, Sapp PC, Horvitz HR, Landers JE, Brown RH Jr (2009) Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. Science 323:1205–1208

Kwon I, Kato M, Xiang S, Wu L, Theodoropoulos P, Mirzaei H, Han T, Xie S, Corden JL, McKnight SL (2013) Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. Cell 155:1049–1060

Lagier-Tourenne C, Polymenidou M, Hutt KR, Vu AQ, Baughn M, Huelga SC, Clutario KM, Ling SC, Liang

TY, Mazur C, Wancewicz E, Kim AS, Watt A, Freier S, Hicks GG, Donohue JP, Shiue L, Bennett CF, Ravits J, Cleveland DW, Yeo GW (2012) Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. Nat Neurosci 15:1488–1497

Lashley T, Rohrer JD, Bandopadhyay R, Fry C, Ahmed Z, Isaacs AM, Brelstaff JH, Borroni B, Warren JD, Troakes C, King A, Al-Saraj S, Newcombe J, Quinn N, Ostergaard K, Schroder HD, Bojsen-Moller-M, Braendgaard H, Fox NC, Rossor MN, Lees AJ, Holton JL, Revesz T (2011) A comparative clinical, pathological, biochemical and genetic study of fused in sarcoma proteinopathies. Brain 134:2548–2564

Lee BJ, Cansizoglu AE, Suel KE, Louis TH, Zhang Z, Chook YM (2006) Rules for nuclear localization sequence recognition by karyopherin beta 2. Cell 126:543–558

Leemann-Zakaryan RP, Pahlich S, Grossenbacher D, Gehring H (2011) Tyrosine phosphorylation in the C-Terminal Nuclear Localization and Retention Signal (C-NLS) of the EWS protein. Sarcoma 2011:218483

Lerga A, Hallier M, Delva L, Orvain C, Gallais I, Marie J, Moreau-Gachelin F (2001) Identification of an RNA binding specificity for the potential splicing factor TLS. J Biol Chem 276:6807–6816

Liu X, Niu C, Ren J, Zhang J, Xie X, Zhu H, Feng W, Gong W (2013) The RRM domain of human fused in sarcoma protein reveals a non-canonical nucleic acid binding site. Biochim Biophys Acta 1832:375–385

Mackenzie IR, Rademakers R, Neumann M (2010a) TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. Lancet Neurol 9:995–1007

Mackenzie IR, Neumann M, Bigio EH, Cairns NJ, Alafuzoff I, Kril J, Kovacs GG, Ghetti B, Halliday G, Holm IE, Ince PG, Kamphorst W, Revesz T, Rozemuller AJ, Kumar-Singh S, Akiyama H, Baborie A, Spina S, Dickson DW, Trojanowski JQ, Mann DM (2010b) Nomenclature and nosology for neuropathologic subtypes of frontotemporal lobar degeneration: an update. Acta Neuropathol 119:1–4

Mackenzie IR, Ansorge O, Strong M, Bilbao J, Zinman L, Ang LC, Baker M, Stewart H, Eisen A, Rademakers R, Neumann M (2011) Pathological heterogeneity in amyotrophic lateral sclerosis with FUS mutations: two distinct patterns correlating with disease severity and mutation. Acta Neuropathol 122:87–98

Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y, Kinoshita Y, Kamada M, Nodera H, Suzuki H, Komure O, Matsuura S, Kobatake K, Morimoto N, Abe K, Suzuki N, Aoki M, Kawata A, Hirai T, Kato T, Ogasawara K, Hirano A, Takumi T, Kusaka H, Hagiwara K, Kaji R, Kawakami H (2010) Mutations of optineurin in amyotrophic lateral sclerosis. Nature 465:223–226

Murakami T, Qamar S, Lin JQ, Schierle GS, Rees E, Miyashita A, Costa AR, Dodd RB, Chan FT, Michel CH, Kronenberg-Versteeg D, Li Y, Yang SP, Wakutani Y, Meadows W, Ferry RR, Dong L, Tartaglia GG, Favrin G, Lin WL, Dickson DW, Zhen M, Ron D, Schmitt-Ulms G, Fraser PE, Shneider NA, Holt C, Vendruscolo M, Kaminski CF, St George-Hyslop P (2015) ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function. Neuron 88:678–690

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, McCluskey LF, Miller BL, Masliah E, Mackenzie IR, Feldman H, Feiden W, Kretzschmar HA, Trojanowski JQ, Lee VM (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 314:130–133

Neumann M, Rademakers R, Roeber S, Baker M, Kretzschmar HA, Mackenzie IR (2009) A new subtype of frontotemporal lobar degeneration with FUS pathology. Brain 132:2922–2931

Neumann M, Bentmann E, Dormann D, Jawaid A, DeJesus-Hernandez M, Ansorge O, Roeber S, Kretzschmar HA, Munoz DG, Kusaka H, Yokota O, Ang LC, Bilbao J, Rademakers R, Haass C, Mackenzie IR (2011) FET proteins TAF15 and EWS are selective markers that distinguish FTLD with FUS pathology from amyotrophic lateral sclerosis with FUS mutations. Brain 134:2595–2609

Neumann M, Valori CF, Ansorge O, Kretzschmar HA, Munoz DG, Kusaka H, Yokota O, Ishihara K, Ang LC, Bilbao JM, Mackenzie IR (2012) Transportin 1 accumulates specifically with FET proteins but no other transportin cargos in FTLD-FUS and is absent in FUS inclusions in ALS with FUS mutations. Acta Neuropathol 124:705–716

Nomura T, Watanabe S, Kaneko K, Yamanaka K, Nukina N, Furukawa Y (2013) Intranuclear aggregation of mutant FUS/TLS as a molecular pathomechanism of amyotrophic lateral sclerosis. J Biol Chem 289:1192–1202

Page T, Gitcho MA, Mosaheb S, Carter D, Chakraverty S, Perry RH, Bigio EH, Gearing M, Ferrer I, Goate AM, Cairns NJ, Thorpe JR (2011) FUS immunogold labeling TEM analysis of the neuronal cytoplasmic inclusions of neuronal intermediate filament inclusion disease: a frontotemporal lobar degeneration with FUS proteinopathy. J Mol Neurosci 45:409–421

Pokrishevsky E, Grad LI, Yousefi M, Wang J, Mackenzie IR, Cashman NR (2012) Aberrant localization of FUS and TDP43 is associated with misfolding of SOD1 in amyotrophic lateral sclerosis. PLoS ONE 7:e35050

Polymenidou M, Cleveland DW (2011) The seeds of neurodegeneration: prion-like spreading in ALS. Cell 147:498–508

Riggi N, Cironi L, Suva ML, Stamenkovic I (2007) Sarcomas: genetics, signalling, and cellular origins. Part 1: The fellowship of TET. J Pathol 213:4–20

Robberecht W, Philips T (2013) The changing scene of amyotrophic lateral sclerosis. Nat Rev Neurosci 14:248–264

Rogelj B, Easton LE, Bogu GK, Stanton LW, Rot G, Curk T, Zupan B, Sugimoto Y, Modic M, Haberman N, Tollervey J, Fujii R, Takumi T, Shaw CE, Ule J (2012) Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. Sci Rep 2:603

Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, Donaldson D, Goto J, O'Regan JP, Deng HX et al (1993) Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 362:59–62

Sama RR, Ward CL, Kaushansky LJ, Lemay N, Ishigaki S, Urano F, Bosco DA (2013) FUS/TLS assembles into stress granules and is a prosurvival factor during hyperosmolar stress. J Cell Physiol 228:2222–2231

Scherzinger E, Lurz R, Turmaine M, Mangiarini L, Hollenbach B, Hasenbank R, Bates GP, Davies SW, Lehrach H, Wanker EE (1997) Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo. Cell 90:549–558

Schwartz JC, Wang X, Podell ER, Cech TR (2013) RNA seeds higher-order assembly of FUS protein. Cell Rep 5:918–925

Schwartz JC, Cech TR, Parker RR (2015) Biochemical Properties and Biological Functions of FET Proteins. Annu Rev Biochem 84:355–379

Sharma A, Lyashchenko AK, Lu L, Nasrabady SE, Elmaleh M, Mendelsohn M, Nemes A, Tapia JC, Mentis GZ, Shneider NA (2016) ALS-associated mutant FUS induces selective motor neuron degeneration through toxic gain of function. Nat Commun 7:10465

Shaw DJ, Morse R, Todd AG, Eggleton P, Lorson CL, Young PJ (2009) Identification of a tripartite import signal in the Ewing Sarcoma protein (EWS). Biochem Biophys Res Commun 390:1197–1201

Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, Ackerley S, Durnall JC, Williams KL, Buratti E, Baralle F, de Belleroche J, Mitchell JD, Leigh PN, Al-Chalabi A, Miller CC, Nicholson G, Shaw CE (2008) TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. Science 319:1668–1672

Suarez-Calvet M, Neumann M, Arzberger T, Abou-Ajram C, Funk E, Hartmann H, Edbauer D, Kremmer E, Gobl C, Resch M, Bourgeois B, Madl T, Reber S, Jutzi D, Ruepp MD, Mackenzie IR, Ansorge O, Dormann D, Haass C (2016) Monomethylated and unmethylated FUS exhibit increased binding to Transportin and distinguish FTLD-FUS from ALS-FUS. Acta Neuropathol 131:587–604

Sun Z, Diaz Z, Fang X, Hart MP, Chesi A, Shorter J, Gitler AD (2011) Molecular determinants and genetic modifiers of aggregation and toxicity for the ALS disease protein FUS/TLS. PLoS Biol 9:e1000614

Takeuchi R, Toyoshima Y, Tada M, Shiga A, Tanaka H, Shimohata M, Kimura K, Morita T, Kakita A, Nishizawa M, Takahashi H (2013) Transportin 1 accumulates in FUS inclusions in adult-onset ALS without FUS mutation. Neuropathol Appl Neurobiol 39:580–584

Tan AY, Manley JL (2009) The TET family of proteins: functions and roles in disease. J Mol Cell Biol 1:82–92

Turner BJ, Talbot K (2008) Transgenics, toxicity and therapeutics in rodent models of mutant SOD1-mediated familial ALS. Prog Neurobiol 85:94–134

Van Deerlin VM, Leverenz JB, Bekris LM, Bird TD, Yuan W, Elman LB, Clay D, Wood EM, Chen-Plotkin AS, Martinez-Lage M, Steinbart E, McCluskey L, Grossman M, Neumann M, Wu IL, Yang WS, Kalb R, Galasko DR, Montine TJ, Trojanowski JQ, Lee VM, Schellenberg GD, Yu CE (2008) TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. Lancet Neurol 7:409–416

Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, Sreedharan J, Hu X, Smith B, Ruddy D, Wright P, Ganesalingam J, Williams KL, Tripathi V, Al-Saraj S, Al-Chalabi A, Leigh PN, Blair IP, Nicholson G, de Belleroche J, Gallo JM, Miller CC, Shaw CE (2009) Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science 323:1208–1211

Wang WY, Pan L, Su SC, Quinn EJ, Sasaki M, Jimenez JC, Mackenzie IR, Huang EJ, Tsai LH (2013) Interaction of FUS and HDAC1 regulates DNA damage response and repair in neurons. Nat Neurosci 16:1383–1391

Woulfe J, Gray DA, Mackenzie IR (2010) FUS-immunoreactive intranuclear inclusions in neurodegenerative disease. Brain Pathol 20:589–597

Zhang ZC, Chook YM (2012) Structural and energetic basis of ALS-causing mutations in the atypical proline-tyrosine nuclear localization signal of the Fused in Sarcoma protein (FUS). Proc Natl Acad Sci U S A 109:12017–12021

Zinszner H, Sok J, Immanuel D, Yin Y, Ron D (1997) TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. J Cell Sci 110(Pt 15):1741–1750

Zoghbi HY, Orr HT (2000) Glutamine repeats and neurodegeneration. Annu Rev Neurosci 23:217–247

Structural Changes Fundamental to Gating of the Cystic Fibrosis Transmembrane Conductance Regulator Anion Channel Pore

Paul Linsdell

Abstract

Cystic fibrosis is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR), an epithelial cell anion channel. Potentiator drugs used in the treatment of cystic fibrosis act on the channel to increase overall channel function, by increasing the stability of its open state and/or decreasing the stability of its closed state. The structure of the channel in either the open state or the closed state is not currently known. However, changes in the conformation of the protein as it transitions between these two states have been studied using functional investigation and molecular modeling techniques. This review summarizes our current understanding of the architecture of the transmembrane channel pore that controls the movement of chloride and other small anions, both in the open state and in the closed state. Evidence for different kinds of changes in the conformation of the pore as it transitions between open and closed states is described, as well as the mechanisms by which these conformational changes might be controlled to regulate normal channel gating. The ways that key conformational changes might be targeted by small compounds to influence overall CFTR activity are also discussed. Understanding the changes in pore structure that might be manipulated by such small compounds is key to the development of novel therapeutic strategies for the treatment of cystic fibrosis.

P. Linsdell (✉)

Department of Physiology & Biophysics, Dalhousie University, PO Box 15000, Halifax, NS B3H 4R2, Canada

e-mail: paul.linsdell@dal.ca

1   Introduction: Cystic Fibrosis and CFTR

Cystic fibrosis (CF) is the most common lethal genetic disease among Caucasian populations, in which it affects 1 in 2500–4500 newborns; over 70,000 individuals worldwide are currently living with CF (O'Sullivan and Freedman 2009; Lubamba et al. 2012). Cystic fibrosis is an autosomal recessive disease, caused by mutations in the gene that encodes the cystic fibrosis transmembrane conductance regulator (CFTR) (Riordan et al. 1989). CFTR (ABCC7) is a member of a large family of membrane transport proteins, the ATP-binding cassette (ABC) transporters (Dean et al. 2001; Rees et al. 2009). Most ABC proteins are active transport ATPases that couple unidirectional transmembrane transport of substrate to ATP hydrolysis (Rees et al. 2009; ter Beek et al. 2014). CFTR appears to be a unique exception to this ABC family practice – it acts not as an active pump but as an ion channel that mediates the passive, electrodiffusional movement of Cl^-, HCO_3^- and other small anions across the cell membrane (Gadsby et al. 2006; Hwang and Kirk 2013; Linsdell 2014a). CFTR is widely expressed in the body, predominantly in the apical (luminal) membrane of epithelial cells lining ducts and tubes, where it is involved in transepithelial fluid and electrolyte transport and pH regulation (Frizzell and Hanrahan 2012; Collawn and Matalon 2014). As a result, loss of CFTR function in CF patients results in wide-ranging clinical manifestations including respiratory disease, pancreatic insufficiency, meconium ileus, diabetes, liver disease, and infertility, especially in males (O'Sullivan and Freedman 2009; Lubamba et al. 2012). Currently, pulmonary insufficiency as a result of chronic airway infection and inflammation resulting in bronchiectasis is the cause of most CF-related mortality.

Around 2000 different mutations in CFTR have been associated with CF, although to date only approximately 160 have been shown definitively to cause CF. These CF-associated mutations have been classified into six categories depending on their predominant mechanism of CFTR dysfunction: defective protein production, defective protein processing, defective channel regulation, defective channel conduction, reduced protein synthesis, and reduced protein stability (Lubamba et al. 2012; Wang et al. 2014c). Different pharmacological strategies may be required to overcome the basic defect associated with different mutation categories, leading to the emergence of the concept of "personalized medicine" or mutation-specific drug therapies for CF patients (Lubamba et al. 2012; Ikpa et al. 2014). Because of the inviolable relationship between CFTR dysfunction and CF, most interest is currently focused on identification and development of small molecules that interact directly with the CFTR protein to modulate its function (Lubamba et al. 2012; Bosch and De Boeck 2016). These small molecules are grouped into "correctors" that improve folding and trafficking of mutant CFTR protein to the cell membrane, and "potentiators" that increase the channel function of CFTR present in the cell membrane. Ivacaftor (Vx-770), the first drug approved for the treatment of CF that directly targets the molecular cause of the disease, is a potentiator (Bosch and De Boeck 2016; Wainright 2014).

In terms of potentiator action, an increase in channel function is defined as an increase in the proportion of time the channel spends in the open, conducting state relative to the closed, non-conducting state (referred to as the channel open probability). Physiological regulation of CFTR channel activity is predominantly via cyclic AMP/protein kinase A (PKA) signaling (Frizzell and Hanrahan 2012); PKA-dependent phosphorylation of the CFTR protein is a prerequisite for channel opening (Chong et al. 2013; Hwang and Kirk 2013). In phosphorylated CFTR, channel gating (opening and closing) is controlled by the binding and hydrolysis of cytoplasmic ATP molecules associated with CFTR's ATPase activity (Gadsby et al. 2006; Kirk and Wang 2011; Jih and Hwang 2012) (see below, "CFTR Channel Gating"). However, Vx-770 is thought to act directly on phosphorylated CFTR, independently of ATP-dependent gating, to

increase its open probability (Eckford et al. 2012; Jih and Hwang 2013).

The assumption that Vx-770 and other potentiators act to increase CFTR activity by directly stabilizing the open channel conformation and/or destabilizing the closed channel conformation provides important impetus to understanding the structural changes that take place in the channel as it transitions between open and closed states. The structure of CFTR – either in the open state or the closed state – is not currently known, and so direct information on the nature of these conformational changes is lacking. On the other hand, since CFTR is an ion channel the conformational change between the open and closed states has been observed at the functional level in individual CFTR molecules using single channel patch clamp recording for over 25 years. Even in the absence of a known structure, the detailed functional information now available as a result of extensive electrophysiological investigation of CFTR allows a mechanistic view of the conformational changes associated with CFTR channel gating to be developed. In this review I will summarize the functional evidence for different types of conformational changes associated with opening and closing of the CFTR channel pore, and attempt to frame this emerging functional model in the context of what is known about the structure and function of other ABC proteins and other ion channels.

2 Understanding the Transport Mechanism in CFTR – A Channel Derived from a Pump

Understanding the relationship between structure and function in CFTR – and in particular the conformational changes that control its transport function – is complicated by the apparent incongruity between its presumed structure (as an ABC protein whose relatives are active transporters) and its function (as an ion channel). Presumably at some point during evolutionary history the function of some CFTR-ancestor protein changed sufficiently for it to cross the conceptual divide from pump to channel.

Both active transporter pumps and ion channels form a substrate-translocation pathway that is shielded from the membrane lipid bilayer, effectively increasing the permeability of the selected substrate(s) across the membrane. For most active transporters, including ABC proteins (ter Beek et al. 2014), it is thought that the transmembrane transport pathway must exist in at least two different conformations in which the substrate is alternately exposed to one side of the membrane or the other (Fig. 1). Controlled switching between these two conformations then allows substrate to be translocated across the membrane by a so-called "alternating access" mechanism (Fig. 1). Recently, it has been suggested that some active transporters may undergo more dramatic conformational rearrangements as they transition between the "inward-facing" (substrate transport pathway open to the inside) and "outward-facing" (substrate transport pathway open to the outside) conformations illustrated in Fig. 1. However, whether such an "elevator" mechanism of alternating access (Ryan and Vandenberg 2016) has any relevance to ABC protein-mediated transport is not currently known. Switching between inward-facing and outward-facing conformations may be coupled to a source of energy (such as ATP hydrolysis in the case of ABC proteins and other primary active transporters) or to co- or counter-transport of another substrate (in the case of secondary active transporters) to allow active accumulation of the primary substrate against its transmembrane concentration gradient. Crystal structures of active transport proteins in inward-facing and outward-facing conformations has allowed some insight into the nature and extent of conformational changes associated with transport in primary active transporters (Nyblom et al. 2013; Toyoshima and Cornelius 2013), secondary active transporters (Forrest et al. 2011; Shi 2013; Reithmeier and Moraes 2015; Ryan and Vandenberg 2016) and facilitative transporters (Yan 2015; Quistgaard et al. 2016). In general these structures suggest that widespread conformational rearrangements of the membrane-spanning parts of the protein have to occur in order to transport between one and a handful of
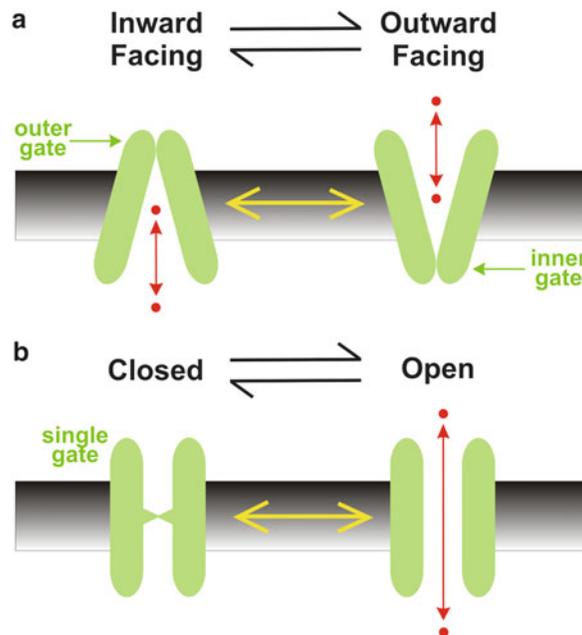
Fig. 1 Conformational changes underlying transmembrane transport mediated by pumps and by channels. (**a**) Membrane pumps are considered as acting by an alternating access model; the substrate (*red*) is made alternately accessible to the intracellular solution (*left*) or to the extracellular solution (*right*) by a global conformational change in the membrane-spanning parts of the pump from inward-facing to outward-facing. This inward-to-outward-facing conformational rearrangement can also be considered as the concurrent opening of an outer gate and closure of an inner gate. For a pump to mediate active transport of substrate across the membrane, it is forbidden for both gates to be open at the same time. More complex schemes incorporating intermediate occluded states (both gates closed) and large rearrangements of the transport pathway have recently been proposed for some active transport proteins (Ryan and Vandenberg 2016). (**b**) Membrane ion channels are considered as acting by the operation of a single gate, that prevents ion flow when closed and allows ion flow when open. The existence of a fully open, continuous substrate transport pathway across the membrane is therefore one crucial functional distinction between pump and channel-mediated transport mechanisms

substrate molecules (Forrest et al. 2011; Nyblom et al. 2013; Shi 2013; Toyoshima and Cornelius 2013; Reithmeier and Moraes 2015; Yan 2015; Quistgaard et al. 2016; Ryan and Vandenberg 2016).

For ion channels, such dramatic conformational rearrangements of the membrane-spanning parts of the protein might not be necessary to control substrate transport. Channel pores are thought to transition almost instantaneously between different conformations that either prevent ion movement ("closed" conformations) or allow very rapid (up to ~10^8 ions s^{-1}) electrodiffusional ion flux ("open" conformations) (Fig. 1). Conceptually, ion flux is thought of as being controlled by a single "gate" in the pore (Fig. 1), the opening and closing of which may be coupled to a distant "sensor" to allow channel

opening and closing to be controlled by factors such as ligand binding or membrane potential (Hille 2001). High-resolution structures of ion channels expected to be in the open state or in the closed state, for ligand-gated channels (Baconguis et al. 2013; Unwin 2013; Sobolevsky 2015), voltage-gated channels (Bagnéris et al. 2015; Catterall and Zheng 2015; Oiki 2015) and other channel types (Zhou and McCammon 2010; Oiki 2015) are consistent with channel opening and closing reflecting a relatively minor structural change in the transmembrane channel pore itself, perhaps controlled by a more major structural rearrangement at the more distant ligand-binding or voltage-sensing parts of the channel protein that is then transmitted over distance to the gate itself (Zhou and McCammon 2010; Baconguis et al. 2013;

Unwin 2013; Bagnéris et al. 2015; Catterall and Zheng 2015; Oiki 2015; Sobolevsky 2015).

Although the transport mechanisms of active transporters and ion channels are traditionally thought of as being very different, CFTR is not alone in its apparent existence at the boundary between the two (Chen and Hwang 2008; Gadsby 2009). For example, different members of the ClC protein family can act as either passive Cl^- channels or as active Cl^-/H^+ exchangers (Chen and Hwang 2008; Gadsby 2009; Accardi 2015), with no obvious structural difference between the two (Accardi 2015). At the same time, channel-like pores may exist (at least transiently) in many types of active transporters (DeFelice and Goswami 2007; Gadsby 2009; Li et al. 2013). It has been pointed out that the difference between channels and pumps really boils down to the number of gates operating on the substrate transport pathway (at least two for a pump, only one for a channel), and how the activity of these gates is coordinated: for a pump, all gates must never be open simultaneously, as must occur in the open state of a channel (Gadsby 2009; Miller 2010) (Fig. 1). This simple discriminator has led to the description of CFTR as a "broken pump" or "degraded transporter", in which one gate has either ceased to function (close) and/or its opening and closing has become uncoupled from the normal gating control mechanism (Gadsby 2009; Jordan et al. 2008; Mornon et al. 2009; Miller 2010; Bai et al. 2011). What vestiges remain of the ancestral (active) transport mechanism – and how much can be learned about CFTR function from the active transport mechanism of other ABCs proteins (and vice versa) – remain open questions.

3 The Structure and Function of the CFTR Anion Channel

The structure of CFTR has been observed directly only at low resolution (Rosenberg et al. 2011) but is expected to be broadly similar to closely-related ABC protein exporters (Hunt et al. 2013). These ABC proteins, including CFTR, share a modular architecture, with two cytoplasmic nucleotide binding domains (NBDs) that bind and hydrolyze ATP, coupled to two membrane-spanning domains (MSDs) that form the substrate transport pathway (Fig. 2). CFTR also has a unique cytoplasmic regulatory domain that is the site of regulation by PKA-dependent phosphorylation (Chong et al. 2013; Hwang and Kirk 2013) (Fig. 3a). The two NBDs form a head-to-tail dimer with two ATP-binding sites at the dimer interface. The NBDs interact with the MSDs via so-called coupling helices. In all ABC proteins, it is presumed that ATP binding and hydrolysis by the NBDs controls the conformation – and thereby the transport function – of the MSDs. For active transporters, this conformational rearrangement is expected to reflect a switch between inward-facing and outward-facing conformations of the MSDs to ensure active transport via an alternating access-type model (Rees et al. 2009; ter Beek et al. 2014) (Figs. 1 and 2c). For CFTR, ATP binding and hydrolysis by the NBDs controls opening and closing of the anion channel pore (see below, "CFTR Channel Gating").

High-resolution structures have been obtained for a number of different ABC exporters (ter Beek et al. 2014; Choudhury et al. 2014; Lee et al. 2014; Kim et al. 2015). Consistent with the prevailing transport models, these structures demonstrate different degrees of NBD association (tightly dimerized, partially dissociated, fully dissociated) and either inward-facing or outward-facing conformations of the MSDs (Fig. 2b) (ter Beek et al. 2014). Atomic homology models for CFTR have been developed based on several of these non-ion channel templates (Mornon et al. 2008, 2009, 2015; Serohijos et al. 2008; Dalton et al. 2012; Norimatsu et al. 2012; Rahman et al. 2013; Corradi et al. 2015), and in some cases molecular dynamics simulations have been used to explore potential conformational rearrangements of the protein (Rahman et al. 2013; Mornon et al. 2015). These homology models are subject to a number of important caveats, including low functional similarity (between the Cl^- ion channel CFTR and templates that are active transporters of large substrates), low sequence
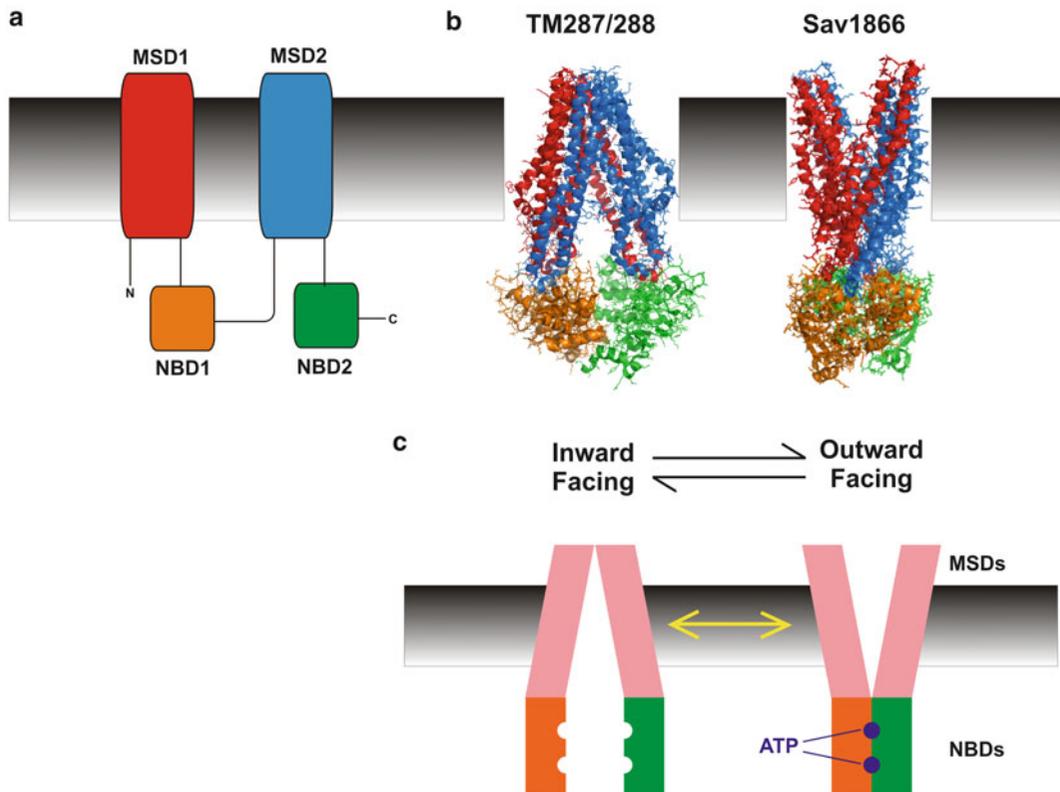
Fig. 2 Architecture and proposed mechanism of ABC proteins. (**a**) Overall modular construction of ABC exporter proteins, comprising two membrane-spanning domains (MSDs) made up of multiple transmembrane α-helices, and two cytoplasmic nucleotide binding domains (NBDs). In this figure these four domains are depicted as being formed by a single polypeptide, as is the case for CFTR, although in some cases the transporter is formed as a dimer of subunits comprising one MSD and one NBD (ter Beek et al. 2014). (**b**) Atomic structures of two ABC proteins illustrating this modular architecture (using the same colour scheme as in (A)): the bacterial export proteins TM287/288 (crystallized with an inward-facing conformation of the MSDs and partially separated NBDs) (Hohl et al. 2012) and Sav1866 (crystallized with an outward-facing conformation of the MSDs and dimerized NBDs) (Dawson and Locher 2007). (**c**) Proposed transport mechanism of ABC exporter proteins. It is assumed that ATP binding to the NBDs promotes NBD dimerization and drives the MSDs into an outward-facing conformation; and that ATP hydrolysis then favours NBD dissociation and drives the MSDs into an inward-facing conformation (Rees et al. 2009; ter Beek et al. 2014)

homology (especially in the MSDs), lack of unique protein regions such as the regulatory domain, and (at least in the case of the earlier models) inconsistencies with known functional properties of the channel (reviewed in Linsdell 2014a). In spite of these qualifications, these models may be able to give valuable insight into the nature and extent of conformational changes that may take place during the transition between open and closed states (Corradi et al. 2015; Mornon et al. 2015) (Fig. 3b and c).

Functional investigation of CFTR using patch clamp recording has been used to develop a working model of the anion channel pore. It is now widely accepted that, in the open state, the channel has a narrow central region that is the main determinant of anion selectivity, flanked by wider inner and outer vestibules that used positively charged amino acid side chains to attract negatively charged ions to the pore (Hwang and Kirk 2013; Linsdell 2014a) (Fig. 3d). Mechanistically important pore-lining amino acids that contribute to the functional properties of each
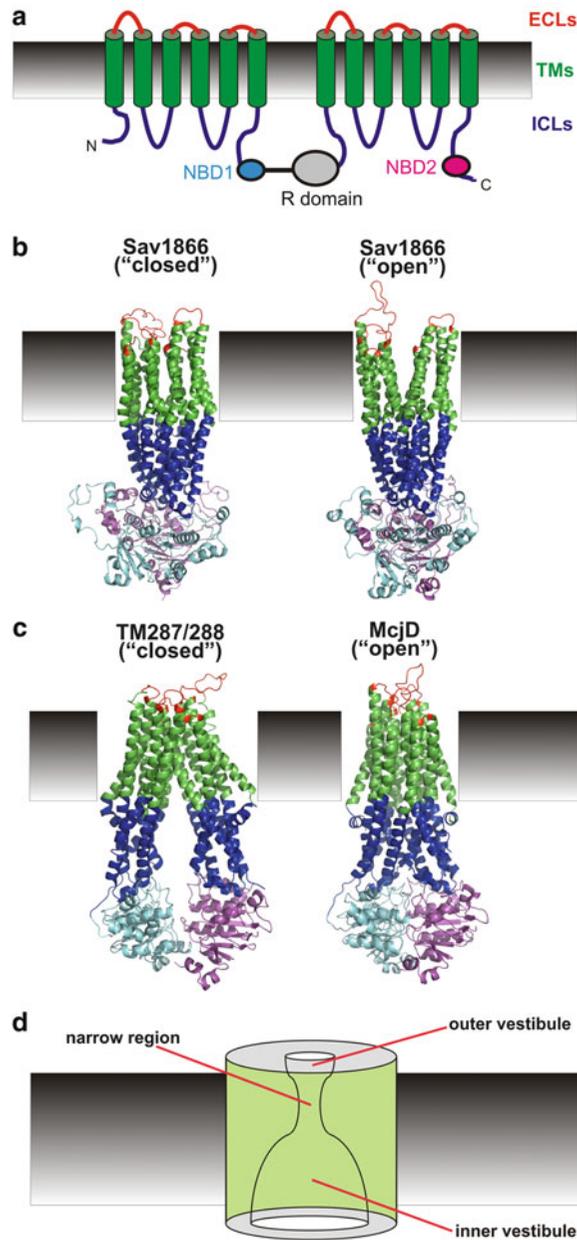
Fig. 3 Different depictions of CFTR channel architecture. (**a**) Schematic two-dimensional representation of CFTR domain structure. The two MSDs are each made up of six transmembrane α-helices (TMs; *green*), connected by short extracellular loops (ECLs; *red*) and longer intracellular loops (ICLs; *blue*). Each MSD is followed by a cytoplasmic NBD (NBD1, *cyan*; NBD2, *magenta*). The two homologous MSD-NBD "halves" of the protein are connected by the unique cytoplasmic regulatory domain (*grey*). (**b, c**) Current atomic models of CFTR, proposed to represent the closed state (*left*) and the open state (*right*) of the channel. (**b**) is based on homology modeling of the bacterial exporter Sav1866 and molecular dynamics simulations (Mornon et al. 2015). (**c**) is based on homology modeling of two different bacterial exporters, TM287/288 and McjD (Corradi et al. 2015). In both (**b**) and (**c**), the approximate extent of the ECLs, TMs, ICLs, and NBDs is indicated by the use of the same colour scheme as in (**a**); the regulatory domain is not included in any of these models. (**d**) Overall architecture of the open channel pore based on functional investigation. As described in the text, the open pore is thought to have a relatively narrow central portion, which is connected to the extracellular solution by a shallow, narrow outer vestibule, and connected to the intracellular solution by a deeper, wider inner vestibule

of these regions of the pore have been identified using extensive structure-function investigations (reviewed by Linsdell 2014a). The membrane-spanning pore is lined by transmembrane (TM) segments TM1, TM6, TM11 and TM12 (El Hiani and Linsdell 2014a), suggesting that the TMs are not arranged symmetrically in the open state (Wang et al. 2014a). The TM6 region plays a particularly important role in determining the functional properties of each of the different regions of the pore (Linsdell 2014a). More recently, it has been shown that the membrane-spanning pore is connected to the cytoplasm by a "lateral tunnel" in one side of the cytoplasmic extension of the MSDs (Corradi et al. 2015; El Hiani and Linsdell 2015; Mornon et al. 2015; El Hiani et al. 2016). Because patch clamp recording gives information on ion flow through the open state of the channel, functional investigation of the closed state(s) of the channel has been more difficult to come by, although such information is important in order to understand the conformational changes in the channel protein during opening and closing transitions.

4 CFTR Channel Gating

Since its introduction in the 1970s, single channel patch clamp recording has offered a rare opportunity to observe functionally a biologically relevant conformational change in a single protein molecule, in its native environment, in real time – namely, an ion channel opening and closing to regulate ion flux while embedded in the cell membrane. As a result, the process of ion channel gating has been studied in great detail for almost all known channel types. Usually a channel molecule is considered to transition almost instantaneously between any number of interconnected open and closed states, resulting in the opening and closing of a single gate within the ion permeation pathway. In fact, this model of gating has held up well in the face of the emergence of realistic ion channel structures. In diverse channel types, specific pore-lining amino acid side-chains may move during channel opening and closing, in order to physically occlude

the pore and/or to form a "hydrophobic seal" that energetically prevents the access of water molecules (and therefore ions), and in this way form a functional "gate" that regulates the movement of ions through the pore (Zhou and McCammon 2010; Baconguis et al. 2013; Unwin 2013; Aryal et al. 2015; Catterall and Zheng 2015; Oiki 2015).

CFTR channel gating is controlled by ATP action at the NBDs (Csanády et al. 2010; Kirk and Wang 2011; Jih and Hwang 2012), as summarized in Fig. 4. As with all ABC proteins, ATP binding promotes dimerization of the two NBDs in a head-to-tail dimer that occludes two ATP molecules in an NBD "sandwich". The stability of this sandwich is such that the NBD dimer separates only after ATP hydrolysis. In CFTR, only one of the two ATP binding sites is catalytically active; ATP remains bound to the other, non-catalytic site during several cycles of NBD dimerization and ATP hydrolysis. ATP binding and NBD dimer formation results in channel opening (Fig. 4). Hydrolysis of ATP at the catalytically active site then causes a transition between two distinct open states, referred to as "O_1" and "O_2" (Fig. 4); the functional properties of these two open states are almost indistinguishable under most conditions (Gunderson and Kopito 1995; Ishihara and Welsh 1997). Following ATP hydrolysis, the NBD dimer is destabilized, and partial dissociation of the NBDs causes channel closure (Fig. 4). Because the ATP sandwich form of the NBDs (as in the O_1 state) is very stable compared to the ADP-bound state following ATP hydrolysis (as in the O_2 state), the majority of CFTR openings are terminated by ATP hydrolysis, resulting in an effective unidirectional "gating cycle" powered by ATP binding and hydrolysis (Csanády et al. 2010) (Fig. 4).

As with other ABC proteins, it is assumed that what goes on at the NBDs in CFTR is somehow transmitted to a change in the conformation of the MSDs. As an ion channel, this conformational change is then reflected in opening and closing of the channel pore (Kirk and Wang 2011; Wang et al. 2014b; Sorum et al. 2015). By analogy with the presumed relationship
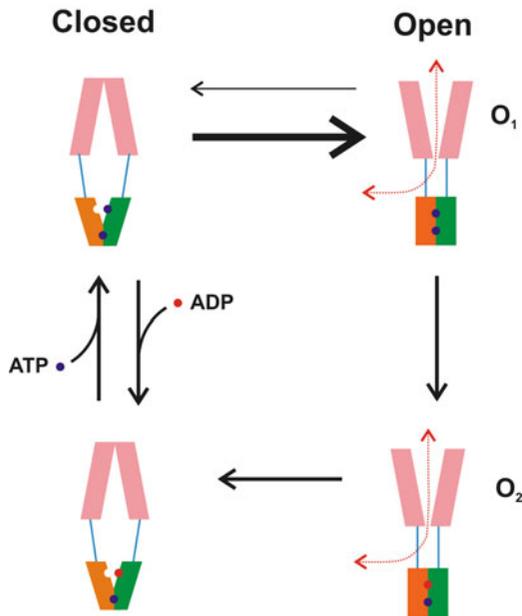
Fig. 4 Proposed CFTR gating mechanism (Csanády et al. 2010; Sorum et al. 2015), modeled after the generic ABC mechanism outlined in Fig. 2c. Binding of ATP to the catalytic site promotes NBD dimerization and channel opening to the O_1 state. Under these conditions, the high stability of the ATP-NBD dimer prevents channel closure until ATP hydrolysis drives the channel into the O_2 state, following which the NBD dimer dissociates and the channel closes. ADP: ATP exchange at the catalytic site then completes the gating cycle. Note that ATP is thought to remain bound to the non-catalytic site for several rounds of the gating cycle, perhaps indicating that only partial NBD dissociation takes place in the closed channel. More complex gating schemes, incorporating multiple ATP hydrolysis steps during a single open burst, have also been proposed (Jih and Hwang 2012)

between the NBDs and the MSDs in related ABC protein exporters – and because NBD dimerization is associated with channel opening, and NBD dissociation with channel closure – it is often supposed that the NBD-dimerized open states will show outwardly-facing MSDs, and the NBD-dissociated closed states will show inwardly-facing MSDs (Gadsby 2009; Mornon et al. 2009; Miller 2010; Kirk and Wang 2011; Hunt et al. 2013; Sorum et al. 2015). However, as discussed below, functional studies of the channel pore itself raise some interesting questions concerning both the validity and the relevance of this widely-held supposition.

5 Inferring the Structural Rearrangements During Gating

5.1 Global Structure of the Pore

Longstanding functional investigation of the CFTR channel pore using permeant and blocking ions has led to the development of a model of the open channel, with a central narrow region, a deep, wide inner vestibule and a shorter, narrower outer vestibule (Hwang and Kirk 2013; Linsdell 2014a) (Fig. 3d). Conceptually this open channel model could be considered similar to an "inward-facing" arrangement of the MSDs, with a permeation pathway that is wide open to the cytoplasm. The structure of the closed channel is more difficult to characterize functionally, but has been probed by substances that can interact with native or introduced sites inside the pore while it is closed, including cysteine-modifying reagents (El Hiani and Linsdell 2014a), metal ions (El Hiani and Linsdell 2014b) and pore-blocking anions (Linsdell 2014b).

Investigation of pore architecture using substituted cysteine accessibility mutagenesis supports this overall functional model of the pore. Sites throughout several TMs have been identified, that can be modified by large cysteine-reactive methanethiosulfonate (MTS) reagents applied to the intracellular side of the membrane, to the extracellular side, or to either side (reviewed in El Hiani and Linsdell 2014a). This side-dependence of modification corresponds approximately to the inner vestibule, the outer vestibule, and the narrow pore region, respectively (Fig. 5a). Channel state-dependent changes in accessibility from either side of the membrane has also been used to infer changes in pore architecture in closed channels (Bai et al. 2010, 2011; El Hiani and Linsdell 2010; Qian et al. 2011; Wang et al. 2011, 2014a; Wang and Linsdell 2012a, b, c; Gao et al. 2013; Zhang and Hwang 2015). Many sites in the inner vestibule, including K95 (TM1), F311 (TM5), V345 (TM6), and S1141 (TM12) have been shown to

remain accessible even to large probes applied to the cytoplasmic side of the membrane in closed channels (El Hiani and Linsdell 2010; Bai et al. 2011; Qian et al. 2011; Wang et al. 2011; Wang and Linsdell 2012a; Zhang and Hwang 2015), indicating that the inner part of the pore remains wide open to the cytoplasm throughout the channel gating cycle. Impermeant, pore-blocking $Pt(NO_2)_4^{2-}$ anions have also been shown to be able to enter from the cytoplasm into the pore in both open and closed channels (apparently with little or no discrimination) to interact with a binding site located in the inner vestibule around the level of K95 (Linsdell 2014a).

Somewhat paradoxically, several studies have found that access from the extracellular solution to the outer vestibule and narrow region of the pore decreases when the channel opens, for example at L102 (Wang and Linsdell 2012b), R334 (Zhang et al. 2005; Wang and Linsdell 2012c), K335 (Beck et al. 2008; Gao and Hwang 2015), and T338 (Wang and Linsdell 2012b, 2012c). These results have been used to suggest that the outer mouth of the pore physically constricts as it opens (Wang and Linsdell 2012b). Most provocatively, it was suggested that the closed channel, being relatively widely accessible to the extracellular solution, showed the closest functional resemblance to an "outward-facing" conformation of the pore (Wang and Linsdell 2012b).

The ability of large cytoplasmic substances to penetrate deep into the pore in both open and closed channels suggests that no functionally relevant gate is capable of closing the pore at its intracellular end (Fig. 5a). This finding could be considered inconsistent with a large global rearrangement of the MSDs during CFTR channel gating, such as that implied by classical alternating access models (Figs. 1a and 2c). Indeed, the most recent atomic homology models feature relatively modest structural rearrangements of the MSDs (Corradi et al. 2015; Mornon et al. 2015) (Fig. 3b), compared to the more overt inward-facing-to-outward-facing transitions of earlier structural models (Mornon et al. 2009). Attempts to correlate one functional conformation of the MSDs (ie. closed, open) as "inward-facing" and the other as "outward-facing" may no longer be helpful in the absence of strong functional data that a large-scale global rearrangement of the kind implied by such terminology is actually associated with channel gating.

5.2   Location of the Channel Gate(s)

As an ion channel, CFTR function is expected to be controlled by a single gate (see above, "Understanding the Transport Mechanism in CFTR – a Channel Derived from a Pump"). The finding that large cytoplasmic substances can penetrate into some parts of the pore inner vestibule when it is closed (El Hiani and Linsdell 2010; Bai et al. 2011; Qian et al. 2011; Wang et al. 2011) indicates that the functionally relevant gate must be located closer to the extracellular end of the pore. Residues located more deeply into the pore, for example Q98 (TM1) and V345 (TM6), have been shown to exhibit state-dependent accessibility to large, cytoplasmically applied MTS reagents (Wang and Linsdell 2012a), suggesting that these substances must wait for the channel to open before they can access this region of the pore from the inside. However, smaller Cl^- ions could, in theory, penetrate more deeply into closed channels, suggesting that the gate could be located closer to the extracellular end of the pore. To address this issue, cysteine-reactive $Au(CN)_2^-$ ions, which are more similar in size to Cl^- and which are highly permeant in the CFTR channel (Smith et al. 1999; Gong et al. 2002), have been used as probes of the putative gate. State-dependent accessibility of intracellular and/or extracellular $Au(CN)_2^-$ to cysteines substituted for Q98 (Wang and Linsdell 2012a), L102 (Wang and Linsdell 2012b), T338 (El Hiani and Linsdell 2010; Wang and Linsdell 2012b; Gao and Hwang 2015) and I344 (Wang and Linsdell 2012a; but see Gao and Hwang 2015) suggest that the gate is located in the narrowest region of the channel pore between T338 and I344 in TM6 (Gao and Hwang 2015) and aligned parts of other TMs (El Hiani and Linsdell 2014a) (Fig. 5).
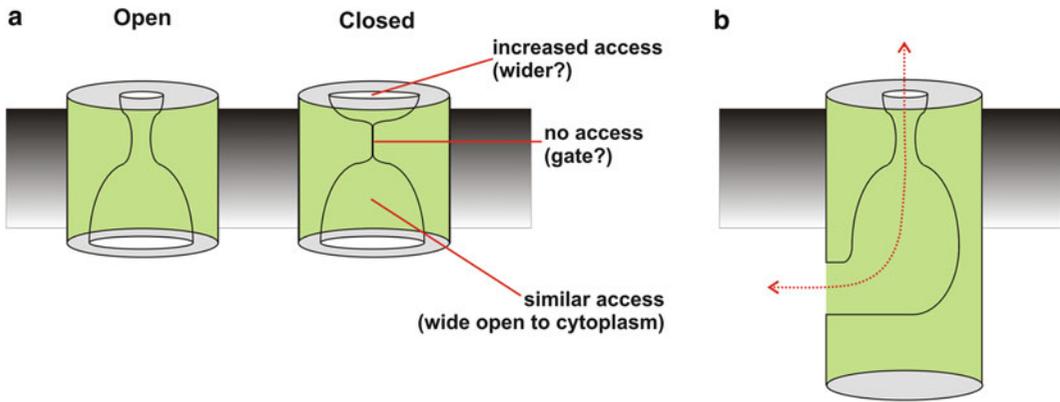
Fig. 5 Proposed functional architecture of the CFTR pore region in open and closed conformations. (**a**) The overall shape of the open channel pore is as depicted in Fig. 3d. As described in the text, closure of the channel has been proposed to be associated with a number of changes in the overall shape of the pore: increased access from the extracellular solution to the outer vestibule may indicate a widening of this part of the pore, loss of access to residues in the narrow part of the pore may indicate the approximate location of the channel gate, and relatively similar access from the intracellular solution to the inner vestibule may indicate that this region of the pore remains wide open to the cytoplasm throughout the gating cycle. (**b**) More recent work suggests that the transmembrane pore is connected to the cytoplasm by a wide lateral portal formed by the ICLs, suggesting that Cl⁻ ions do not follow a central pathway through the entire CFTR protein (as indicated by the *red dotted line*)

Because CFTR descended from the active transporter ABC family, it has been proposed that it may still have some remnant of a "vestigial" second gate that has become degraded or uncoupled from ATP-dependent gating as the protein evolved from an active transporter to an ion channel (Jordan et al. 2008; Gadsby 2009; Mornon et al. 2009; Miller 2010; Bai et al. 2011; Wang and Linsdell 2012b). The outwardly-facing structures of ABC protein exporters (Dawson and Locher 2007; Choudhury et al. 2014; ter Beek et al. 2014) – as well as CFTR atomic homology models that are based on these outwardly-facing active transporter templates (Mornon et al. 2009, 2015; Dalton et al. 2012; Corradi et al. 2015) – have MSDs that are closed at the intracellular side, at the so-called tetrahelix bundle close to the coupling helices that form the link between the intracellular loops (ICLs) and NBDs. In CFTR, this tetrahelix bundle is formed by ICL2 (the cytoplasmic extension of TM3 and TM4) and ICL4 (cytoplasmic extension of TM9 and TM10). Because this structure appears to seal off the intracellular end of the permeation pathway, it has been proposed as the location of a vestigial gate (Mornon et al. 2009; Billet et al. 2013). However, although this tetrahelix bundle does appear to be involved in ATP/NBD-dependent gating of CFTR (Wang et al. 2014b) – consistent with its being in the transmission pathway between the NBDs and the pore – this part of the protein does not contribute to the Cl⁻ permeation pathway itself (El Hiani and Linsdell 2012a) and so cannot form a gate for substrate movement. It now appears likely that Cl⁻ ions by-pass this part of the protein completely, passing between the cytoplasm and the central pore via lateral portals between the ICLs, in particular that between the cytoplasmic extensions to TM4 and TM6 (Corradi et al. 2015; El Hiani and Linsdell 2015; Mornon et al. 2015; El Hiani et al. 2016). In this scenario, CFTR may have evolved from a pump into a channel not because one of its gates became dysfunctional, but because a small side-entrance to the transport pathway allowed one gate to be short-circuited, at least by small Cl⁻ ions (Fig. 5b). In effect, the "broken pump" CFTR is broken because it has a crack in it, near the intracellular end of the transport pathway.

5.3 What Kinds of Movements Do the MSDs Undergo During Gating?

In order for the gate to open and close in response to ATP action at the NBDs, the MSDs must undergo movements that convey the opening transition signal to the gate and to generate its rearrangement to allow or prevent the passage of ions. Different kinds of movements of different TM segments that make up the MSDs have been inferred from experiments that discern gating-dependent changes in the accessibility of individual amino acids, gating-dependent changes in the relative proximity of different amino acids, and the effects of constraining the relative movement of different TMs on gating. The following summarizes the evidence for movements in different regions of the channel pore, beginning at its extracellular end. These different proposed movements are also depicted in Fig. 6.

(i) Extracellular Loops The twelve TMs that make up the MSDs are connected by six short extracellular loops (ECLs). Disulfide bonds formed between cysteine residues introduced into ECL1 and ECL4 impair channel activity, which was taken to suggest that relative movement of these two ECLs is important for normal channel function (Broadbent et al. 2014). Charged residues in these two ECLs may also form state-dependent salt bridges that stabilize either the open state or the closed state (Cui et al. 2014). Sites of engineered disulfide crosslink formation between these ECLs included ECL4 residue R899 (Broadbent et al. 2014), which is also involved in stimulation of channel activity by extracellular anions (Li et al. 2012; Broadbent et al. 2015). It is possible, therefore, that extracellular anions may regulate CFTR activity via changes in the interactions between different ECLs.
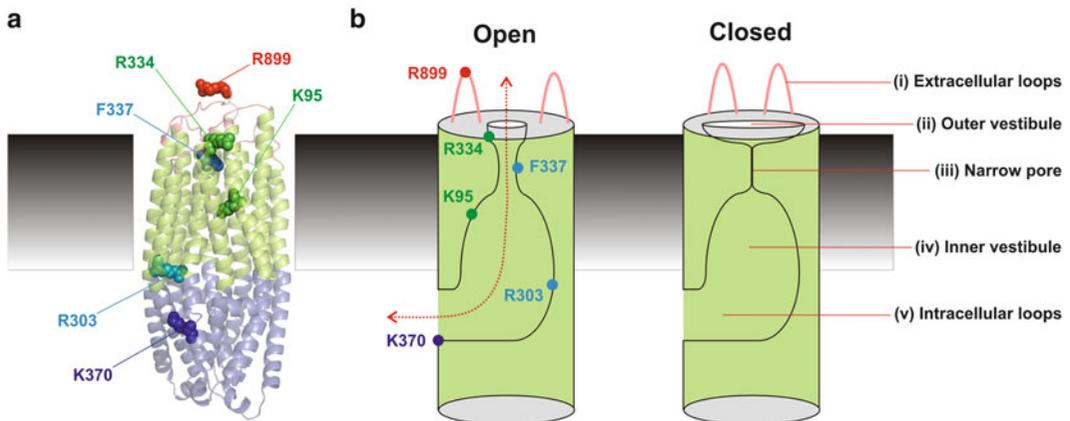
Fig. 6 Summary of proposed conformational changes taking place during opening and closing of the CFTR pore. (**a**) Location of example key amino acid side chains in different parts of the CFTR pore region – namely the extracellular loops (R899), outer vestibule (R334), narrow pore region (F337), inner vestibule (K95) and cytoplasmic entrance (R303, K370) – depicted in a current atomic homology model of the open CFTR channel, based on the bacterial ABC protein McjD (Corradi et al. 2015). For clarity, the NBDs have been removed from this model, and only the MSDs (TMs, ECLs and ICLs) that form the Cl$^-$ transport pathway are depicted. The functional importance of these amino acids in Cl$^-$ permeation is described elsewhere (Linsdell 2014a;

Broadbent et al. 2015; El Hiani and Linsdell 2015). (**b**) Cartoon summary of proposed conformational changes in these different regions of the pore. As described in the text, evidence suggests that channel closure is associated with (i) a relative movement of different extracellular loops, (ii) widening of the outer vestibule, together with relative lateral movement of different TMs, (iii) closure of the narrow pore region, again possibly together with relative lateral movement of different TMs, (iv) rotation, convergence, and separation of different TMs that has little overall impact on accessibility from the cytoplasm to the inner vestibule, and (v) little functional change in the pathway connecting the inner vestibule to the cytoplasm

(ii) Outer Vestibule As described above, state-dependent accessibility experiments have been used to suggest that the outer vestibule of the pore physically constricts during channel opening (Wang and Linsdell 2012b). At the same time, state-dependent disulfide cross-link formation between the outer ends of TM6 and TM11 suggested that these two TMs undergo a relative translational movement during channel opening and closing, with TM6 moving "down" (towards the cytoplasm) relative to TM11 during channel opening, and "up" when the channel closes (Wang and Linsdell 2012c). Crosslinking and molecular dynamics studies have been used to suggest that a salt-bridge forms between TM3 and TM6 in closed channels, and that this salt bridge breaks when the channel opens (Rahman et al. 2013).

(iii) Narrow Pore Relative "downward" translational movement of TM6 could also contribute to the reduced accessibility from the outside, and increased accessibility from the inside, of T338 in the narrow region of the pore in open channels (Wang and Linsdell 2012b, c). It has also been proposed that movement of the bulky side-chain of F337 may regulate water and ion flow through the channel (Corradi et al. 2015), raising interest in this residue as perhaps contributing to the channel gate.

(iv) Inner Vestibule Although much of the inner vestibule appears to be accessible to the cytoplasm both in open channels and in closed channels, a number of different kinds of molecular movements have been proposed to occur in this region during channel gating. Based on state-dependent access by cytoplasmic MTS reagents, both TM6 and TM12 have been suggested to show a rotational movement of up to $100°$ during opening and closing (Bai et al. 2010, 2011), although it has recently been proposed that these reported state-dependent accessibility changes could reflect changes in TM helix-helix interactions rather than rotation of individual helices (Corradi et al. 2015). The apparently increased accessibility of some sites in the inner vestibule to very large MTS adducts in the closed state has been used to suggest that the inner

mouth of the pore is, somewhat paradoxically, wider in closed channels than it is in open channels (Bai et al. 2011; Zhang and Hwang 2015). Most surprisingly, it has been proposed that TM1 undergoes a large-scale movement from being non-pore-lining in closed channels to pore-lining in open channels (Gao et al. 2013), although other work shows pore-lining parts of TM1 (including K95) as being accessible to the cytoplasm in both open and closed channels (Wang et al. 2011; Wang and Linsdell 2012a). "Removal" of TM1 from the pore in closed channels seems difficult to reconcile with the finding that K95 is able to interact electrostatically with pore-blocking anions both in open channels and in closed channels (Linsdell 2014b). Recently, metal ion bridges formed between cysteine side-chains introduced into different TMs were used to show that, during channel opening, TM6 and TM12 move apart, while TM1 and TM12 move closer together (El Hiani and Linsdell 2014b). While this study did not rule out either rotational or translational movement of these TMs during the gating transition, it did suggest that lateral separation and convergence of different TMs is the key conformational change that is actually required to open and close the channel: metal bridges that held TM6 and TM12 together stabilized the closed channel, whereas a metal bridge that held TM1 together with TM12 stabilized the channel in the open state (El Hiani and Linsdell 2014b).

(v) Intracellular Loops Although the cytoplasmic extensions of several TMs have now been shown to contribute to the permeation pathway (El Hiani and Linsdell 2015; El Hiani et al. 2016), conformational changes in this part of the pore during channel gating have not yet been reported. Since large substances are supposed to be able to access the inner vestibule of the pore not only in open channels but also in closed channels, large changes in the structure of the most cytoplasmic part of the permeation pathway might not be anticipated. Nevertheless, there is a significant discrepancy between the apparent large functional diameter of the inner pore (that can allow large organic blocking anions and MTS reagents to penetrate far into

the inner vestibule) (Linsdell 2014c) and the narrow diameter of the lateral tunnels suggested by homology models (barely large enough to accommodate a Cl^- ion) (Corradi et al. 2015; Mornon et al. 2015). This discrepancy might suggest that large changes in the conformation of the inner entrance(s) to the pore are possible.

5.4 Effect of Mutations on Channel Gating

Mutations that disrupt the function of the channel gate – or that disrupt the propagation of the gating signal from the NBDs to the gate – might be expected to disrupt normal ATP-dependent channel gating. Identification and characterization of such mutations could help to identify key conformational rearrangements that are required for the gate to open and close, as well as the molecular mechanism of gate function. For example, mutation of amino acids that contribute to the channel gate may be able to alter gate function and thereby change channel activity. The suggestion that the narrow region of the CFTR channel pore comprises the channel gate (Corradi et al. 2015; Gao and Hwang 2015) (Figs. 5 and 6) might therefore suggest that mutations in this region could interfere with gate function and alter channel gating. Consistent with this idea, mutations that reduce amino acid side-chain volume at the putative narrow region (F337A, F337S) not only disrupt anion selectivity (Linsdell et al. 2000) and conductance (Linsdell 2001) but also increase channel open probability and ATP-independent channel opening (Wei et al. 2016). Furthermore, a nearby mutation in TM1 (L102C) resulted in frequent "flickery" closures during an open burst (perhaps suggesting an unstable open state), and modification of this cysteine side-chain with a positively-charged MTS reagent then caused the channel to become almost permanently locked in the open state, even after the removal of ATP (Gao et al. 2013). Although this result appears consistent with disabling of the channel gate, similar results were obtained for other pore-lining side-chains located further from the narrow region (Bai et al. 2010) (see below), and so the molecular mechanism of this change in gating awaits complete characterization.

Mutations and amino acid side-chain modifications in other parts of the MSDs also affect gating, most likely by affecting the conformational transition that is propagated from the NBDs to the gate in the pore. Both mutations (Cotten et al. 1996; Seibert et al. 1996a, b; Wang et al. 2010, 2014b; Sorum et al. 2015) and MTS-modifications (El Hiani and Linsdell 2012a) within the ICLs significantly affect gating, consistent with these regions being involved in transmission of the gating signal from the NBDs to the pore (Kirk and Wang 2011; Sorum et al. 2015). Most significantly, K978C (in the cytoplasmic extension to TM9 in ICL3) dramatically increases ATP-independent channel opening (Wang et al. 2010; Okeyo et al. 2014), which has been interpreted as reflecting an important role for this residue in the allosteric mechanism that couples conformational changes in the NBDs with those at the gate (Kirk and Wang 2011; Okeyo et al. 2014). Similar effects on nucleotide-independent gating were observed in P355A, at the cytoplasmic end of TM6 (Wei et al. 2014), consistent with a spreading conformational rearrangement. Modification of pore-lining I344C and M348C in TM6 by a positively-charged MTS reagent also greatly increased ATP-independent channel opening (Bai et al. 2010), consistent with movement of TM6 being required for channel closure.

6 How Can the Conformation of the Channel Pore Be Manipulated to Control CFTR Activity?

Since loss of CFTR function results in CF, the ultimate goal of understanding CFTR channel gating at the structural level is to be able to exert control over CFTR activity. Small drugs that can stabilize the open state and/or destabilize the closed state of the channel can help to

maximize CFTR function and so bring relief to CF patients. This is no longer a theoretical goal, but the presumed mechanism of action of Vx-770 (Eckford et al. 2012; Jih and Hwang, 2013), a drug that is currently used successfully to treat some CF patients (Bosch and De Boeck 2016; Wainwright 2014). The binding site – and hence the precise molecular mechanism of action – for Vx-770 is unknown. As described in the preceding paragraphs ("Inferring the Structural Rearrangements During Gating"), widespread conformational changes likely take place as the gating signal is propagated from the NBDs to the gate in the pore, and theoretically a drug that could bind at any site along the propagation axis could act to alter the relative stability of the open and closed states and therefore influence the overall activity of the channel.

One class of drugs that have been shown to affect CFTR channel gating is channel blockers. Binding of $Pt(NO_2)_4^{2-}$ ions in the inner vestibule blocks Cl^- permeation through the open channel (Gong and Linsdell 2003; Linsdell 2014b), and has also been suggested to interfere both with channel opening and channel closing (Linsdell 2014b). The binding site for $Pt(NO_2)_4^{2-}$ includes K95 in TM1 (Zhou et al. 2010; Linsdell 2014b) and is also close to I344 in TM6 (El Hiani and Linsdell 2012b; Linsdell 2014b, 2015). This suggests that binding of $Pt(NO_2)_4^{2-}$ to this site interferes with the conformational rearrangements of the inner vestibule that are required both for channel opening and for channel closing (El Hiani and Linsdell 2014b). A large and diverse range of anionic substances act via a similar molecular mechanism to block Cl^- permeation (Linsdell 2014c), although in most cases the ability of these blockers also to affect channel gating has not been investigated. Blocker binding inside the pore has frequently been shown to prevent the closure of other ion channel types by what is often referred to as a "foot-in-the-door" mechanism, in which the channel gate cannot close until the blocker has exited the pore (Armstrong 1971; Neher and Steinbach 1978; Vilin et al. 2013). In theory, a substance that bound to this site within the inner vestibule of the CFTR pore to prevent channel closure without disrupting Cl^- flow could

hold the channel in the open state and therefore act as a potentiator.

Other small monovalent anions besides Cl^- can carry current through CFTR (Linsdell and Hanrahan 1998), and recent experiments showing differences in open probability (Sorum et al. 2015; Yeh et al. 2015) and open burst length (Sorum et al. 2015) when different anions carry current suggest that binding of permeant anions inside the pore can also hold the channel open. It was suggested that anions that bind more tightly inside the pore result in a more stable open state (Sorum et al. 2015), suggesting that permeant anions may also act as a foot in the CFTR door (or gate). As with many blockers (see above), binding of permeant anions inside the pore is thought to involve the positively charged side chain of K95 in the inner vestibule (Rubaiy and Linsdell 2015; Linsdell 2016), perhaps consistent with the inner vestibule being a site at which substances could bind to increase CFTR channel open probability.

7 Conclusions and Future Perspectives

As an ABC protein, CFTR is expected to show large-scale structural changes in its MSDs associated with ATP action at the NBDs (Figs. 1a and 2), which has been interpreted as implying global rearrangements between inward-facing and outward-facing MSDs, and likewise between closed and open channel pores (Fig. 3). On the other hand, as an ion channel it is feasible that local, rather than global, changes in MSD structure are sufficient to open and close the pore (Fig. 1b). As summarized in Fig. 6, there is now a great deal of functional evidence supporting different kinds of structural rearrangements during opening and closing of the pore. In general, these studies have emphasized local, rather than global, changes in pore conformation during gating. While this may be in part due to the nature of the experimental investigations in question, some aspects of the closed-to-open conformational transition – such as the proposed existence of a local gate in the pore, and the continuous access

from the cytoplasm to a wide-open inner vestibule to the pore in both open and closed states (Fig. 6) – seem somewhat incongruous with the most literal interpretation of the inward-facing-to-outward-facing transition depicted in Figs. 1 and 2. This raises the question whether the ion channel/ABC protein CFTR undergoes a more subtle variant of an inward-facing-to-outward-facing transition; or even if these two distinct conformations of the MSDs exist in CFTR in a manifestation that would be recognizable to more traditional ABC protein experimentalists. Since inward-facing and outward-facing are structural concepts, the answer to these questions must await more direct structural information on the CFTR protein. Nevertheless, known anomalies between CFTR and other ABC proteins – such as proposed uncoupling (or loose coupling) between the NBDs and the MSDs (Jih and Hwang 2012, 2013), the presence of the unique cytoplasmic regulatory domain, and the presence of the lateral cytoplasmic entrance to the pore that might short-circuit any conformational changes taking place close to the NBD: MSD boundary – mean that it is wise to keep an open mind on the overall conformation of CFTR's MSDs in different functional states.

Improved understanding of the conformational changes associated with CFTR channel gating offers an opportunity to manipulate these changes to maximize CFTR activity, which we now know can be beneficial therapeutically for CF patients. However, many challenges still exist in order to seize this opportunity. Figure 6 summarizes many different proposed (local) changes in the shape and orientation of different pore-forming parts of the protein during channel gating, including the lateral convergence, separation, rotation, and translation of different TMs and ECLs, local changes in pore diameter, gate opening and closing, and even parts of the pore for which evidence for functionally important conformational changes is currently lacking. Of all these potential changes in conformation, which are crucial for opening and closing to take place, and which might be considered coincidental? Which can be manipulated by small compounds to alter the relative stability of the

open and closed states depicted in cartoon form in Figs. 1b and 6b? How does the gate actually operate on the molecular level, to allow or prevent the passage of Cl^- and other small anions? How is the conformation of the pore controlled by the NBDs, and how is the signal initiated by ATP action at the NBDs transmitted to the gate? Which steps along this putative transmission pathway from the NBDs to the gate offer the best opportunity for pharmacological manipulation, and how is this affected by different classes of CF mutations? The answers to these questions will be key to translating the basic science outlined in this review into tangible benefit for CF patients.

References

Accardi A (2015) Structure and gating of CLC channels and exchangers. J Physiol 593:4129–4138

Armstrong CM (1971) Interaction of tetraethylammonium ion derivatives with the potassium channels of giant axons. J Gen Physiol 58:413–437

Aryal P, Sansom MSP, Tucker SJ (2015) Hydrophobic gating in ion channels. J Mol Biol 427:121–130

Baconguis I, Hattori M, Gouaux E (2013) Unanticipated parallels in architecture and mechanism between ATP-gated P2X receptors and acid sensing ion channels. Curr Opin Struct Biol 23:277–284

Bagnéris C, Naylor CE, McCusker EC, Wallace BA (2015) Structural model of the open-closed-inactivated cycle of prokaryotic voltage-gated sodium channels. J Gen Physiol 145:5–16

Bai Y, Li M, Hwang T-C (2010) Dual roles of the sixth transmembrane segment of the CFTR chloride channel in gating and permeation. J Gen Physiol 136:293–309

Bai Y, Li M, Hwang T-C (2011) Structural basis for the channel function of a degraded ABC transporter, CFTR (ABCC7). J Gen Physiol 138:495–507

Beck EJ, Yang Y, Yaemsiri S, Raghuram V (2008) Conformational changes in a pore-lining helix coupled to cystic fibrosis transmembrane conductance regulator channel gating. J Biol Chem 283:4957–4966

Billet A, Mornon J-P, Jollivet M, Lehn P, Callebaut I, Becq F (2013) CFTR: effect of ICL2 and ICL4 amino acids in close spatial proximity on the current properties of the channel. J Cyst Fibros 12:737–745

Bosch B, De Boeck K (2016) Searching for a cure for cystic fibrosis. A 25-year quest in a nutshell. Eur J Pediatr 175:1–8

Broadbent SD, Wang W, Linsdell P (2014) Interaction between two extracellular loops influences the activity of the cystic fibrosis transmembrane conductance regulator chloride channel. Biochem Cell Biol 92:390–396

Broadbent SD, Ramjeesingh M, Bear CE, Argent BE, Linsdell P, Gray MA (2015) The cystic fibrosis transmembrane conductance regulator is an extracellular chloride sensor. Pflugers Arch 467:1783–1794

Catterall WA, Zheng N (2015) Deciphering voltage-gated Na^+ and Ca^{2+} channels by studying prokaryotic ancestors. Trends Biochem Sci 40:526–534

Chen T-Y, Hwang T-C (2008) CLC-0 and CFTR: chloride channels evolved from transporters. Physiol Rev 88:351–387

Chong PA, Kota P, Dokholyan NV, Forman-Kay JD (2013) Dynamics intrinsic to cystic fibrosis transmembrane conductance regulator function and stability. Cold Spring Harb Perspect Med 3:a009522

Choudhury HG, Tong Z, Mathavan I, Li Y, Iwata S, Zirah S, Rebuffat S, van Veen HW, Beis K (2014) Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state. Proc Natl Acad Sci U S A 111:9145–9150

Collawn JF, Matalon S (2014) CFTR and lung homeostasis. Am J Physiol 307:L917–L923

Corradi V, Vergani P, Tieleman DP (2015) Cystic fibrosis transmembrane conductance regulator (CFTR). Closed and open state channel models. J Biol Chem 290:22891–22906

Cotten JF, Ostedgaard LS, Carson MR, Welsh MJ (1996) Effect of cystic fibrosis-associated mutations in the fourth intracellular loop of cystic fibrosis transmembrane conductance regulator. J Biol Chem 271:21279–21284

Csanády L, Vergani P, Gadsby DC (2010) Strict coupling between CFTR's catalytic cycle and gating of its Cl^- ion pore revealed by distributions of open channel burst durations. Proc Natl Acad Sci U S A 107:1241–1246

Cui G, Rahman KS, Infield DT, Kuang C, Prince CZ, McCarty NA (2014) Three charged amino acids in extracellular loop 1 are involved in maintaining the outer pore architecture of CFTR. J Gen Physiol 144:159–179

Dalton J, Kalid O, Schushan M, Ben-Tal N, Villà-Freixa J (2012) New model of cystic fibrosis transmembrane conductance regulator proposes active channel-like conformation. J Chem Inf Model 52:1842–1853

Dawson RPJ, Locher KP (2007) Structure of the multidrug ABC transporter Sav 1866 from Staphylococcus aureus in complex with AMP-PNP. FEBS Lett 581:935–938

Dean M, Rzhetsky A, Alikmets R (2001) The human ATP-binding cassette (ABC) transporter superfamily. Genome Res 11:1156–1166

DeFelice LJ, Goswami T (2007) Transporters as channels. Annu Rev Physiol 69:87–112

Eckford PDW, Li C, Ramjeesingh M, Bear CE (2012) Cystic fibrosis transmembrane conductance regulator (CFTR) potentiator VX-770 (ivacaftor) opens the defective channel gate of mutant CFTR in a phosphorylation-dependent but ATP-independent manner. J Biol Chem 287:36639–36649

El Hiani Y, Linsdell P (2010) Changes in accessibility of cytoplasmic substances to the pore associated with activation of the cystic fibrosis transmembrane conductance regulator chloride channel. J Biol Chem 285:32126–32140

El Hiani Y, Linsdell P (2012a) Role of the juxtamembrane region of cytoplasmic loop 3 in the gating and conductance of the cystic fibrosis transmembrane conductance regulator chloride channel. Biochemistry 51:3971–3981

El Hiani Y, Linsdell P (2012b) Tuning of CFTR chloride channel function by location of positive charges within the pore. Biophys J 103:1719–1726

El Hiani Y, Linsdell P (2014a) Conformational changes opening and closing the CFTR chloride channel: insights from cysteine scanning mutagenesis. Biochem Cell Biol 92:481–488

El Hiani Y, Linsdell P (2014b) Metal bridges illuminate transmembrane domain movements during gating of the cystic fibrosis transmembrane conductance regulator chloride channel. J Biol Chem 289:28149–28159

El Hiani Y, Linsdell P (2015) Functional architecture of the cytoplasmic entrance to the cystic fibrosis transmembrane conductance regulator chloride channel pore. J Biol Chem 290:15855–15865

El Hiani Y, Negoda A, Linsdell P (2016) Cytoplasmic pathway followed by chloride ions to enter the CFTR channel pore. Cell Mol Life Sci 73:1917–1925

Forrest LR, Krämer R, Ziegler C (2011) The structural basis of secondary active transport mechanisms. Biochim Biophys Acta 1807:167–188

Frizzell RA, Hanrahan JW (2012) Physiology of epithelial chloride and fluid secretion. Cold Spring Harb Perspect Med 2:a009563

Gadsby DC (2009) Ion channels versus ion pumps: the principal difference, in principle. Nat Rev Mol Cell Biol 10:344–352

Gadsby DC, Vergani P, Csanády L (2006) The ABC protein turned chloride channel whose failure causes cystic fibrosis. Nature 440:477–483

Gao X, Hwang T-C (2015) Localizing a gate in CFTR. Proc Natl Acad Sci U S A 112:2461–2466

Gao X, Bai Y, Hwang T-C (2013) Cysteine scanning of CFTR's first transmembrane segment reveals its plausible roles in gating and permeation. Biophys J 104:786–797

Gong X, Linsdell P (2003) Mutation-induced blocker permeability and multiion block of the CFTR chloride channel pore. J Gen Physiol 122:673–687

Gong X, Burbridge SM, Cowley EA, Linsdell P (2002) Molecular determinants of $Au(CN)_2^-$ binding and permeability within the cystic fibrosis transmembrane conductance regulator Cl^- channel pore. J Physiol 540:39–47

Gunderson KL, Kopito RR (1995) Conformational states of CFTR associated with channel gating: the role of ATP binding and hydrolysis. Cell 82:231–239

Hille B (2001) Ion channels of excitable membranes. Sinauer, Sunderland

Hohl M, Briand C, Grütter MG, Seeger MA (2012) Crystal structure of a heterodimeric ABC transporter in its inward-facing conformation. Nat Struct Mol Biol 19:395–402

Hunt JF, Wang C, Ford RC (2013) Cystic fibrosis transmembrane conductance regulator (ABCC7) structure. Cold Spring Harb Perspect Med 3:a009514

Hwang T-C, Kirk KL (2013) The CFTR ion channel: gating, regulation, and anion permeation. Cold Spring Harb Perspect Med 3:a009498

Ikpa PT, Bijvelds MJ, de Jonge HR (2014) Cystic fibrosis: towards personalized therapies. Int J Biochem Cell Biol 52:192–200

Ishihara H, Welsh MJ (1997) Block by MOPS reveals a conformation change in the CFTR pore produced by ATP hydrolysis. Am J Physiol 273:C1278–C1289

Jih K-Y, Hwang T-C (2012) Nonequilibrium gating of CFTR on an equilibrium theme. Physiology 27:351–361

Jih K-Y, Hwang T-C (2013) Vx-770 potentiates CFTR function by promoting decoupling between the gating cycle and ATP hydrolysis cycle. Proc Natl Acad Sci U S A 110:4404–4409

Jordan IK, Kota KC, Cui G, Thompson CH, McCarty NA (2008) Evolutionary and functional divergence between the cystic fibrosis transmembrane conductance regulator and related ATP-binding cassette transporters. Proc Natl Acad Sci U S A 105:18865–18870

Kim J, Wu S, Tomasiak TM, Mergel C, Winter MB, Stiller SB, Robles-Colmanares Y, Stroud RM, Tampé R, Craik CS, Cheng Y (2015) Subnanometre-resolution electron cryomicroscopy structure of a heterodimeric ABC exporter. Nature 517:396–400

Kirk KL, Wang W (2011) A unified view of cystic fibrosis transmembrane conductance regulator (CFTR) gating: combining the allosterism of a ligand-gated channel with the enzymatic activity of an ATP-binding cassette (ABC) transporter. J Biol Chem 286:12813–12819

Lee JY, Yang JG, Zhitnitsky D, Lewinson O, Rees DC (2014) Structural basis for heavy metal detoxification by an Atm1-type ABC exporter. Science 343:1133–1136

Li M-S, Cowley EA, Linsdell P (2012) Pseudohalide anions reveal a novel extracellular site for potentiators to increase CFTR function. Br J Pharmacol 167:1062–1075

Li J, Shaikh SA, Enkavi G, Wen P-C, Huang Z, Tajkhorshid E (2013) Transient formation of water-conducting states in membrane transporters. Proc Natl Acad Sci U S A 110:7696–7701

Linsdell P (2001) Relationship between anion binding and anion permeability revealed by mutagenesis within the cystic fibrosis transmembrane conductance regulator chloride channel pore. J Physiol 531:51–66

Linsdell P (2014a) Functional architecture of the CFTR chloride channel. Mol Membr Biol 31:1–16

Linsdell P (2014b) State-dependent blocker interactions with the CFTR chloride channel: implications for gating the pore. Pflugers Arch 466:2243–2255

Linsdell P (2014c) Cystic fibrosis transmembrane conductance regulator chloride channel blockers: pharmacological, biophysical and physiological relevance. World J Biol Chem 5:26–39

Linsdell P (2015) Interactions between permeant and blocking anions inside the CFTR chloride channel pore. Biochim Biophys Acta 1848:1573–1590

Linsdell P (2016) Anion conductance selectivity mechanism of the CFTR chloride channel. Biochim Biophys Acta 1858:740–747

Linsdell P, Hanrahan JW (1998) Adenosine triphosphate-dependent asymmetry of anion permeation in the cystic fibrosis transmembrane conductance regulator chloride channel. J Gen Physiol 111:601–614

Linsdell P, Evagelidis A, Hanrahan JW (2000) Molecular determinants of anion selectivity in the cystic fibrosis transmembrane conductance regulator chloride channel pore. Biophys J 78:2973–2982

Lubamba B, Dhooghe B, Noel S, Leal T (2012) Cystic fibrosis: insight into CFTR pathophysiology and pharmacotherapy. Clin Biochem 45:1132–1144

Miller C (2010) CFTR: break a pump, make a channel. Proc Natl Acad Sci U S A 107:959–960

Mornon J-P, Lehn P, Callebaut I (2008) Atomic model of human cystic fibrosis transmembrane conductance regulator: membrane-spanning domains and coupling interfaces. Cell Mol Life Sci 65:2594–2612

Mornon J-P, Lehn P, Callebaut I (2009) Molecular models of the open and closed states of the whole human CFTR protein. Cell Mol Life Sci 66:3469–3486

Mornon J-P, Hoffmann B, Jonic S, Lehn P, Callebaut I (2015) Full-open and closed CFTR channels, with lateral tunnels from the cytoplasm and an alternative position of the F508 region, as revealed by molecular dynamics. Cell Mol Life Sci 72:1377–1403

Neher E, Steinbach JH (1978) Local anaesthetics transiently block currents through single acetylcholine-receptor channels. J Physiol 277:153–176

Norimatsu Y, Ivetac A, Alexander C, Kirkham J, O'Donnell N, Dawson DC, Sansom MS (2012) Cystic fibrosis transmembrane conductance regulator: a molecular model defines the architecture of the anion conduction path and locates a "bottleneck" in the pore. Biochemistry 51:2199–2212

Nyblom M, Poulsen H, Gourdon P, Reinhard L, Andersson M, Lindahl E, Fedosova N, Nissen P (2013) Crystal structure of Na^+, K^+-ATPase in the Na^+-bound state. Science 342:123–127

Oiki S (2015) Channel function reconstitution and re-animation: a single-channel strategy in the postcrystal age. J Physiol 593:2553–2573

Okeyo G, Wang W, Wei S, Kirk KL (2014) Converting nonhydrolyzable nucleotides to strong cystic fibrosis transmembrane conductance regulator (CFTR) agonists by gain of function (GOF) mutations. J Biol Chem 288:17122–17133

O'Sullivan BP, Freedman SD (2009) Cystic fibrosis. Lancet 373:1891–1904

Qian F, El Hiani Y, Linsdell P (2011) Functional arrangement of the 12th transmembrane region in the CFTR chloride channel pore based on functional investigation of a cysteine-less CFTR variant. Pflugers Arch 462:559–571

Quistgaard EM, Löw C, Guettou F, Norlund P (2016) Understanding transport by the major facilitator superfamily (MFS): structures pave the way. Nat Rev Mol Cell Biol 17:123–132

Rahman KS, Cui G, Harvey SC, McCarty NA (2013) Modeling the conformational changes underlying channel opening in CFTR. PLoS ONE 8:e74574

Rees DC, Johnson E, Lewinson O (2009) ABC transporters: the power to change. Nat Rev Mol Cell Biol 10:218–227

Reithmeier RAF, Moraes TF (2015) Solute transporters keep on rockin. Nat Struct Mol Biol 22:752–754

Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou J-L, Drumm ML, Iannuzzi MC, Collins FS, Tsui L-C (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science 245:1066–1073

Rosenberg MF, O'Ryan LP, Hughes G, Zhao Z, Aleksandrov LA, Riordan JR, Ford RC (2011) The cystic fibrosis transmembrane conductance regulator (CFTR). Three-dimensional structure and localization of a channel gate. J Biol Chem 286:42647–42654

Rubaiy HN, Linsdell P (2015) Location of a permeant anion binding site in the cystic fibrosis transmembrane conductance regulator chloride channel pore. J Physiol Sci 65:233–241

Ryan RM, Vandenberg RJ (2016) Elevating the alternating-access model. Nat Struct Mol Biol 23:187–189

Seibert FS, Linsdell P, Loo TW, Hanrahan JW, Clarke DM, Riordan JR (1996a) Disease-associated mutations in the fourth cytoplasmic loop of cystic fibrosis transmembrane conductance regulator compromise biosynthetic processing and chloride channel activity. J Biol Chem 271:15139–15145

Seibert FS, Linsdell P, Loo TW, Hanrahan JW, Riordan JR, Clarke DM (1996b) Cytoplasmic loop three of cystic fibrosis transmembrane conductance regulator contributes to regulation of chloride channel activity. J Biol Chem 271:27493–27499

Serohijos AWR, Hegedüs T, Aleksandrov AA, He L, Cui L, Dokholyan NV, Riordan JR (2008) Phenylalanine-508 mediates a cytoplasmic-membrane domain contact in the CFTR 3D crystal structure crucial to assembly and channel function. Proc Natl Acad Sci U S A 105:3256–3261

Shi Y (2013) Common folds and transport mechanisms of secondary active transporters. Annu Rev Biophys 42:51–72

Smith SS, Steinle ED, Meyerhoff ME, Dawson DC (1999) Cystic fibrosis transmembrane conductance regulator. Physical basis for lyotropic anion selectivity patterns. J Gen Physiol 114:799–818

Sobolevsky AI (2015) Structure and gating of tetrameric glutamate receptors. J Physiol 593:29–38

Sorum B, Czégé D, Csanády L (2015) Timing of CFTR pore opening and structure of its transition state. Cell 163:724–733

ter Beek J, Guskov A, Slotboom DJ (2014) Structural diversity of ABC transporters. J Gen Physiol 143:419–435

Toyoshima C, Cornelius F (2013) New crystal structures of PII-type ATPases: excitement continues. Curr Opin Struct Biol 23:507–514

Unwin N (2013) Nicotinic acetylcholine receptor and the structural basis of neuromuscular transmission: insights from *Torpedo* postsynaptic membranes. Q Rev Biophys 4:283–322

Vilin YY, Nunez J-J, Kim RY, Dake GR, Kurata HT (2013) Paradoxical activation of an inwardly rectifying potassium channel mutant by spermine: "(B)locking" open the bundle crossing gate. Mol Pharmacol 84:572–581

Wainwright CE (2014) Ivacaftor for patients with cystic fibrosis. Expert Rev Respir Med 8:533–538

Wang W, Linsdell P (2012a) Conformational change opening the CFTR chloride channel pore coupled to ATP-dependent gating. Biochim Biophys Acta 1818:851–860

Wang W, Linsdell P (2012b) Alternating access to the transmembrane domain of the ATP-binding cassette protein cystic fibrosis transmembrane conductance regulator (ABCC7). J Biol Chem 287:10156–10165

Wang W, Linsdell P (2012c) Relative movements of transmembrane regions at the outer mouth of the cystic fibrosis transmembrane conductance regulator channel pore during channel gating. J Biol Chem 287:32136–32146

Wang W, Wu J, Bernard K, Li G, Wang G, Bevensee MO, Kirk KL (2010) ATP-independent CFTR channel gating and allosteric modulation by phosphorylation. Proc Natl Acad Sci U S A 107:3888–3893

Wang W, El Hiani Y, Linsdell P (2011) Alignment of transmembrane regions in the cystic fibrosis transmembrane conductance regulator chloride channel pore. J Gen Physiol 138:165–178

Wang W, El Hiani Y, Rubaiy HN, Linsdell P (2014a) Relative contribution of different transmembrane segments to the CFTR chloride channel pore. Pflugers Arch 466:477–490

Wang W, Roessler BC, Kirk KL (2014b) An electrostatic interaction at the tetrahelix bundle promotes phosphorylation-dependent cystic fibrosis transmembrane conductance regulator (CFTR) channel opening. J Biol Chem 289:30364–30378

Wang Y, Wrennall JA, Cai Z, Li H, Sheppard DN (2014c) Understanding how cystic fibrosis mutations disrupt CFTR function: from single molecules to animal models. Int J Biochem Cell Biol 52:47–57

Wei S, Roessler BC, Chauvet S, Guo J, Hartman JL, Kirk KL (2014) Conserved allosteric hot spots in the transmembrane domains of cystic fibrosis transmembrane conductance regulator (CFTR) channels and multidrug resistance protein (MRP) pumps. J Biol Chem 289:19942–19957

Wei S, Roessler BC, Icyuz M, Chauvet S, Tao B, Hartman JL, Kirk KL (2016) Long-range coupling between the extracellular gates and the intracellular ATP binding domains of multidrug resistance protein pumps and cystic fibrosis transmembrane conductance regulator channels. FASEB J 30:1247–1262

Yan N (2015) Structural biology of the major facilitator superfamily transporters. Annu Rev Biophys 44:257–283

Yeh H-I, Yeh J-T, Hwang T-C (2015) Modulation of CFTR gating by permeant ions. J Gen Physiol 145:47–60

Zhang J, Hwang T-C (2015) The fifth transmembrane segment of cystic fibrosis transmembrane conductance regulator contributes to its anion permeation pathway. Biochemistry 54:3839–3850

Zhang Z-R, Song B, McCarty NA (2005) State-dependent chemical reactivity of an engineered cysteine reveals conformational changes in the outer vestibule of the cystic fibrosis transmembrane conductance regulator. J Biol Chem 280:41997–42003

Zhou H-X, McCammon JA (2010) The gates of ion channels and enzymes. Trends Biochem Sci 35:179–185

Zhou J-J, Li M-S, Qi J, Linsdell P (2010) Regulation of conductance by the number of fixed positive charges in the intracellular vestibule of the CFTR chloride channel pore. J Gen Physiol 135:229–245

Dual Roles for Epithelial Splicing Regulatory Proteins 1 (ESRP1) and 2 (ESRP2) in Cancer Progression

Akira Hayakawa, Masao Saitoh, and Keiji Miyazawa

Abstract

Epithelial splicing regulatory protein 1 (ESRP1) and 2 (ESRP2) are members of the hnRNP family of RNA binding proteins that regulate alternative splicing events associated with epithelial phenotypes. These proteins play crucial roles during organogenesis, including craniofacial and epidermal development as well as branching morphogenesis in the lungs and salivary glands. Recent reports have also addressed their roles during cancer progression. Expression of ESRP proteins is low in normal epithelium but upregulated in carcinoma *in situ* and advanced carcinomas. Intriguingly, they are downregulated in invasive fronts. The plastic nature of ESRP expression suggests dual roles for them in cancer progression. Consistently, it has been shown that ESRPs suppress motility and anchorage-independent growth of cancer cells while supporting cell survival by enhancing resistance to reactive oxygen species. Regulatory circuits that fine-tune *ESRP* gene expression have recently emerged. Here, we summarize recent findings on the molecular mechanisms by which ESRPs exert positive as well as negative effects on cancer progression.

Keywords

Epithelial splicing regulatory protein • Alternative splicing • Cell motility • δ-crystallin enhancer binding protein • CD44

_____

A. Hayakawa, M. Saitoh, and K. Miyazawa (✉)
Department of Biochemistry, Interdisciplinary Graduate
School of Medicine, University of Yamanashi,
Yamanashi 408-3898, Japan
e-mail: kmiyazawa@yamanashi.ac.jp

Abbreviations

δEF1	δ-crystallin enhancer binding protein
EMT	epithelial-mesenchymal transition
ESRP	epithelial splicing regulatory protein
FGF	fibroblast growth factor
FGFR	fibroblast growth factor receptor

hnRNP	heterogenous nuclear ribonucleoprotein
HNSCC	head and neck squamous carcinoma
RRM	RNA-recognition motif
SIP1	Smad interacting protein 1.

1 ESRP1 and ESRP2: Splicing Regulatory Proteins Specifically Expressed in Epithelial Cells

Alternative splicing generates mRNA variants that encode protein isoforms with diverse or even opposite functions from pre-mRNAs transcribed from a single gene (David and Manley 2010; Braunschweig et al. 2013; Warzecha and Carstens 2012; Kaida et al. 2012). The basic machinery for pre-mRNA splicing represents the spliceosome, a large complex composed of snRNPs (U1, U2, U4, U5, and U6) and additional 150–200 proteins. Recruitment of the spliceosome to alternative splice sites is tightly regulated by RNA-binding proteins associated with *cis*-regulatory elements in pre-mRNAs, thus ensuring cell-type specific formation of mRNA variants. RNA-binding proteins that regulate alternative splicing events are classified into two families, serine/arginine-rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs). In general, SR proteins interact with exonic or intronic splicing enhancers to promote exon inclusion, whereas hnRNPs interact with exonic or intronic splicing suppressors to promote exon skipping (Warzecha and Carstens 2012; Kaida et al. 2012). The balance between these positive and negative regulators determines the extent of exon inclusion in target pre-mRNAs.

Fibroblast growth factor receptors (FGFRs) are proteins with two isoforms, namely epithelial FGFR-IIIb and mesenchymal FGFR-IIIc. These isoforms have distinct ligand binding properties but share a common intracellular signaling domain. They are produced by cell-type specific alternative splicing (Eswarakumar et al. 2005). The transition between FGFR isoforms affects the behavior of cells by altering their sensitivity to various fibroblast growth factor (FGF) ligands

present in surrounding tissue microenvironments. One such example involves isoform switching of FGFRs during the epithelial-mesenchymal transition (EMT), a process by which epithelial cells lose their polarity and acquire motile and invasive phenotypes (Kalluri and Weinberg 2009). Epithelial cells are usually insensitive to FGF2 that is abundant in tumor tissues because they only express the FGFR-IIIb isoform. In cells that undergo EMT, FGFR-IIIb is downregulated while FGFR-IIIc is upregulated to confer sensitivity to FGF2, thus inducing even more aggressive phenotypes in the cells in the presence of FGF2 (enhanced EMT) (Shirakihara et al. 2011).

By genome-wide cDNA expression screening, Warzecha et al identified two RNA binding proteins belonging to the hnRNP family, RBM35A and RBM35B, as regulators of alternative splicing of the *FGFR2* pre-mRNA (Warzecha et al. 2009a). Expression of RBM35A and 35B is well correlated to that of the epithelial FGFR2-IIIb isoform in various cell lines. In addition, *in situ* hybridization analysis of whole postnatal and adult mice revealed that their expression is epithelium-specific in various tissues and organs. They thus renamed RBM35A and B as epithelial splicing regulatory protein 1 (ESRP1) and 2 (ESRP2). Subsequently ESRPs were shown to bind preferentially to UGG-rich repeats and regulate epithelial specific splicing of a diverse array of target pre-mRNAs (Warzecha et al. 2009b; Warzecha et al. 2010; Dittmar et al. 2012), suggesting that ESRPs function to maintain the epithelial phenotypes of cells. Notably, some of the RNA splice variants regulated by ESRPs have been implicated in regulating cytoskeleton reorganization and cell adhesion (Warzecha et al. 2009b).

2 ESRP Expression Is Upregulated in Cancer Cells

The mechanisms that restrict expression of ESRPs in epithelial cells remain to be fully understood. δ-crystallin enhancer binding protein (δEF1, also called zinc finger E–box binding

homeobox 1, ZEB1), Smad interacting protein 1 (SIP1, also called ZEB2), and Snail, transcriptional repressors that are expressed in cells with mesenchymal phenotypes, were shown to inhibit ESRP expression (Horiguchi et al. 2012; Reinke et al. 2012). Notably, δEF1 and SIP1 directly interact with the promoter regions of the *ESRP* genes as revealed by chromatin-immunoprecipitation assays (Horiguchi et al. 2012). By contrast, Grhl2 (grainyhead-like-2), a transcription factor expressed in epithelial cells, was reported to upregulate *ESRP1* expression in breast cancer cells (Xiang et al. 2012). This upregulation may not be due to a direct effect of Grhl2 because Grhl2 and δEF1 mutually repress each other (Cieply et al. 2013; Werner et al. 2013), suggesting that Grhl2 induces ESRP1 expression by downregulating δEF1 expression.

Determining ESRP expression in human patient specimens would give valuable information on the role of ESRPs in pathological processes. In normal human pancreas, ESRP1 is only weakly expressed in pancreatic ductal cells but it is abundantly expressed in well-to-moderately differentiated adenocarcinoma although downregulated in poorly differentiated adenocarcinoma (Ueda et al. 2014). In normal human oral squamous epithelium, ESRP1 and 2 are weakly expressed in the basal layer (Ishii et al. 2014). In carcinoma *in situ* and advanced carcinomas, they are highly expressed. Importantly, ESRPs disappear from invasive fronts while they are re-expressed in cells that have metastasized to a lymph node. These findings indicate that the expression of ESRPs is plastic during cancer progression.

3 ESRPs Negatively Regulate Cell Motility Through Multiple Mechanisms

Downregulation of ESRPs in invasive fronts suggests negative regulatory roles for ESRPs in cancer invasion and metastasis. Consistently, ESRPs have been shown to suppress cell motility *in vitro*: Knockdown of ESRP1 or ESRP2

promotes cell motility in human mammary epithelial cells (Warzecha et al. 2010), pancreatic cancer cells (Ueda et al. 2014), head and neck squamous cell carcinoma (HNSCC) cells (Ishii et al. 2014), and clear-cell renal cell carcinoma cells (Mizutani et al. 2015). Conversely, the motility and invasiveness of pancreatic cancer cells are attenuated following overexpression of ESRP1 (Ueda et al. 2014).

Lu et al showed that ESRP1 is involved in the alternative splicing of Exo70, which produces two alternatively spliced protein isoforms: epithelially expressed Exo70-E and mesenchymally expressed Exo70-M (Lu et al. 2013). During EMT, the repression of Exo70-E coincides with the repression of ESRP1. Knockdown of ESRP1 in MCF7 breast cancer cells increased the expression of Exo70-M, which can interact with a complex of actin-related proteins, Arp2/3, to stimulate actin polymerization, resulting in promotion of cell migration and invasion. However, it remains to be demonstrated if Exo70-M is indispensable for increased cell motility upon knockdown of ESRP1.

Ishii et al found that knockdown of either ESRP1 or ESRP2 enhanced the motility of HNSCC cell lines (SAS, HSC4) to the similar extent (Ishii et al. 2014). They found that ESRP2 knockdown caused cell dissociation accompanied with downregulation of E-cadherin while ESRP1 knockdown triggered formation of long filopodia. Increased cell motility with long filopodia formation by ESRP1 knockdown was attributed to induction of Rac1b, an alternatively spliced isoform of the Rac1 protein. Rac1 is a small G-protein that regulates reorganization of the actin cytoskeleton. Rac1b, often expressed in cancer cells, is a constitutively active form of Rac1 that contains a 19 amino acid residue insertion (Jordan et al. 1999; Schnelzer et al. 2000; Fiegen et al. 2004). Downregulation of E-cadherin as well as increased cell motility caused by ESRP2 knockdown is due to upregulation of EMT-associated transcription factors, including δEF1 and SIP1, that are known to repress E-cadherin (Comijn et al. 2001; Eger

et al. 2005) and a cell motility-inhibitory protein RGS16 (regulator of G-protein signaling 16) (Hoshi et al. 2016). Thus, ESRP1 and 2 suppress cell motility of HNSCC cells via distinct mechanisms (Fig. 1). Intriguingly, δEF1 and SIP1 suppress the expression of ESRP1 and 2 as described above (Horiguchi et al. 2012), indicating that a double negative feedback circuit is formed between δEF1/SIP1 and ESRP2. This regulatory circuit may facilitate the immediate transition of cellular states as well as plastic expression of ESRPs in response to certain stimuli.

The mechanisms by which ESRP2 downregulates δEF1 and SIP1 remain unclear. ESRP2 may downregulate δEF1/SIP1 through alternative splicing events of certain target pre-mRNAs or through a splicing-independent mechanism in HNSCC cells. Intriguingly, ESRP1 was shown to regulate the translation of
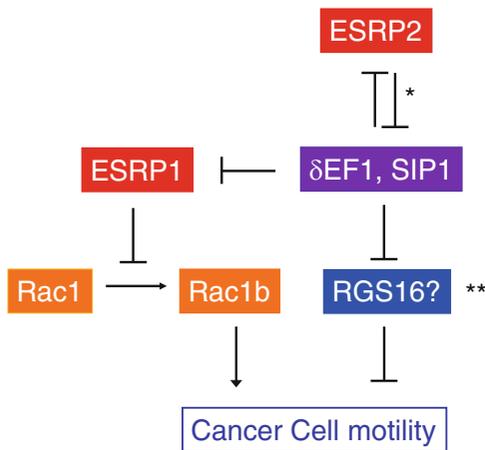
pluripotency-related factors in embryonic stem cells by interacting with the 5'-untranslated region of target mRNAs in the cytoplasm (Fagoonee et al. 2013). In breast cancer cells, ectopic expression of ESRP1 and 2 resulted in higher expression of E-cadherin without affecting the expression levels of δEF1, SIP1, and Snail that downregulate E-cadherin mRNA (CDH1) (Horiguchi et al. 2012). The effects of ESRP2 on δEF1/SIP1 expression may therefore be context-dependent. Recently, Preca et al proposed a model for linking ESRP1 and δEF1 expression that is mediated by isoform switching of CD44 in breast cancer cells (Fig. 2) (Preca et al. 2015). CD44 is a transmembrane glycoprotein that can interact with extracellular matrices, including hyaluronan. Switching between the standard CD44 isoform (CD44s) and variant isoforms (CD44v) is regulated by alternative splicing (Ponta et al. 2003). ESRP1 expression increases CD44v levels in epithelial cells. During EMT, upregulated δEF1 represses ESRP1 expression, triggering the isoform switching of CD44 from CD44v to CD44s. The CD44s isoform, in turn, enhances δEF1 expression by an unknown mechanism, thus forming a positive feedback circuit. In the case of HNSCC cells, however, ESRP2 represses δEF1/SIP1 expression without affecting CD44 isoform switching (Ishii et al. 2014).

Fig. 1 Molecular network that regulates motility of HNSCC cells. ESRP1 and ESRP2 suppress cell motility by downregulating Rac1b and δEF1/SIP1 expression, respectively (Ishii et al. 2014). δEF1/SIP1 in turn repress ESRP1 and ESRP2 expression (Horiguchi et al. 2012), thus forming a double negative feedback circuit between ESRP2 and δEF1/SIP1. The negative effect of ESRP1 on Rac1b expression appears to be due to regulation of the alternative splicing event generating Rac1 and Rac1b mRNAs. *The mechanism by which ESRP2 downregulates δEF1/SIP1 remains to be elucidated. **δEF1 and SIP1 enhance cancer cell motility by downregulating RGS16 in breast cancer cells (Hoshi et al. 2016), but this has not yet been demonstrated in HNSCC cells

Fig. 2 A regulatory circuit comprising ESRP1, δEF1 and CD44s in breast cancer cells. ESRP1 and δEF1 form a double negative feedback loop mediated by CD44s signaling in breast cancer cells (Preca et al. 2015). ESRP1 affects alternative splicing of the CD44 pre-mRNA to upregulate CD44v while downregulate CD44s that enhances expression of δEF1. δEF1 transcriptionally downreguialtes ESRP1, resulting in induction of CD44s that in turn enhances δEF1 expression. *It remains to be elucidated how CD44s, but not CD44v, induces δEF1 expression

Further studies are required to resolve these conflicting observations.

4   Dual Roles of ESRPs on Cancer Progression

Horiguchi et al reported that expression of ESRP1 and ESRP2 is inversely related to cancer malignancy: they are poorly expressed in 'basal-like' subtype of breast cancer cells exhibiting high malignancy, while highly expressed in luminal-type breast cancer cells exhibiting low malignancy (Horiguchi et al. 2012). Basal-like MDA-MB-231 cells that express ESRPs ectopically changed their morphology from spindle to cobble stone–like shape, accompanied by E-cadherin expression, and failed to proliferate efficiently in soft agar (Horiguchi et al. 2012). Leontieva & Ionov also reported that ectopic expression of ESRP1 in LS180 colon cancer cells attenuated anchorage-independent growth *in vitro* and tumorigenic potential *in vivo* (hetero-topic xenograft model) (Leontieva and Ionov 2009). Ueda et al showed that pancreatic cancer cells that overexpress ESRP1 exhibit decreased metastasis to the liver and the lung when they were orthotopically implanted in mice (Ueda et al. 2014). Ueda et al further determined the survival rate in pancreatic ductal adenocarci-noma cases based on their immunohistochemical data. Both overall and the disease-free survival rates of the "ESRP1-high" group were higher than those of the "ESRP1-low" group (Ueda et al. 2014). Consistently, Preca et al reported that pancreatic ductal adenocarcinomas with poor outcome and recurrence expressed low levels of ESRP1 (Preca et al. 2015). These findings based on experiments using ESRP overexpression all suggest that ESRP1 has nega-tive impacts on cancer progression.

Conversely, Yae et al reported that ESRP1-silenced 4T1 breast cancer cells that are orthotopically transplanted exhibit decreased incidence of lung metastasis, thus suggesting a positive role for ESRP1 in metastasis (Yae et al. 2012). The underlying mechanism appears to be as follows: ESRP1 expression results in

increase in expression of the CD44v isoform. CD44v, but not CD44s, interacts with and stabilizes the xCT subunit of a glutamate-cystine transporter, leading to the enhanced uptake of cysteine to facilitate glutathione synthesis (Ishimoto et al. 2011). Thus, ESRP1 supports cancer cell proliferation by increasing levels of cellular glutathione that can serve as an antioxi-dant and confer resistance to reactive oxygen species to cells. They also utilized a public data-base to determine that breast cancer patients expressing high levels of ESRP1 mRNA exhibited a lower rate of overall survival (Yae et al. 2012). A positive role for ESRP1 may be related to the immunohistochemical findings that it is upregulated in carcinoma *in situ* and advanced carcinomas (Ueda et al. 2014; Ishii et al. 2014). However, it has also been reported that CD44s, but not CD44v, confers cells with resistance to cisplatin through activating the phosphoinositide-3-kinase/Akt pathway and CD44s mRNA is enriched in high-grade breast cancers (Brown et al. 2011). The conflicting data may be due to different experimental systems used in these studies.

ESRPs thus appear to have dual roles in can-cer progression depending on the context of microenvironments surrounding cancer cells. In some situations, ESRP expression is favored as it supports cell survival; in other situations, downregulation of ESRPs is favored as this facilitates cell invasion. Therefore, cancer cells that are successful at fine-tuning ESRP expres-sion to adapt surrounding circumstances would be able to undergo further progression.

5   Possible Functional Differences Between ESRP1 and ESRP2

The functional differences between ESRP1 and ESRP2 remain to be clearly elucidated. Both ESRP1 and ESRP2 harbor three RNA-recognition motifs (RRMs) and the amino acid sequences of each motif are well conserved between the two proteins (80–90 % identity) (Fig. 3). Interestingly, a point mutation in
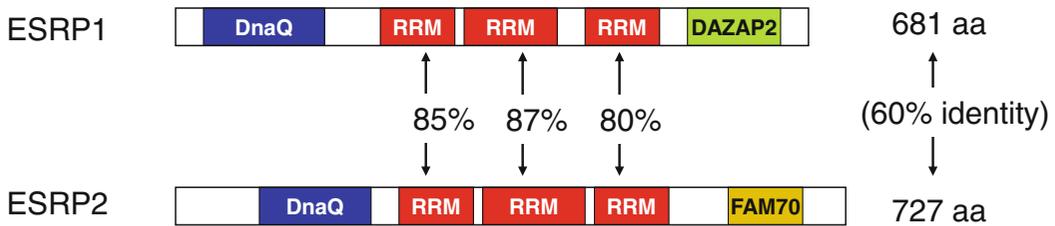
Fig. 3 Schematic structures of ESRP1 and ESRP2. ESRP1 and ESRP2 share the N-terminal DnaQ-like exonuclease domain (DnaQ) and three tandem repeat RNA recognition motifs (RRMs) with a high degree of sequence conservation. The RRM2 and RRM3 of ESRP2 are implicated in target pre-mRNA recognition (Mizutani et al. 2015; Horvath et al. 2013). In the C-terminal region, ESRP1 has a proline-rich region that shares homology with DAZAP2 whereas ESRP2 has a region that shares homology with FAM70. The roles that these domains play in ESRP functions remain unclear

ESRP2 at the second RRM (Arg353Gln) is reported in breast cancers (Horvath et al. 2013). This mutation was shown to impair the ability of ESRP2 to bind to a *cis*-regulatory motif in the *FGFR2* pre-mRNA, suggesting a role for RRM2 in recognizing target pre-mRNAs.

Knockdown of ESRP1, but not ESRP2, altered the isoform switching of CD44 from CD44v to CD44s in HNSCC cells, which suggests distinct roles for both ESRPs (Ishii et al. 2014). However, in normal mouse mammary gland epithelial (NMuMG) cells that lack ESRP1 expression, ESRP2 knockdown resulted in the isoform switching of CD44 (Horiguchi et al. 2012). Thus the target preference of both ESRPs is likely to be context-dependent. It can be affected by endogenous expression levels of ESRPs and target pre-mRNAs. Alternatively, it may be regulated by posttranslational modification of ESRP proteins. Recently, Mizutani et al reported that the splicing activity of ESRP2 is enhanced by K27-linked polyubiquitination of its RRM2 and RRM3 by a ubiquitin ligase Arkadia (Mizutani et al. 2015).

There is also evidence for distinct functions for both proteins *in vivo*. Germline knockout of *Esrp1* resulted in abnormal craniofacial development and neonatal lethality (Bebee et al. 2015). Intriguingly, high expression of *Esrp2* but not *Esrp1* is observed in mouse liver, and similarly in adult human livers for the corresponding human orthologs. *Esrp2*-null mice display an increased number of diploid as well as tetraploid hepatocytes with smaller sizes, suggesting that mouse ESRP2 plays a role in postnatal liver development (Bhate et al. 2015). However, double knockout mouse embryos exhibit more severe phenotypes, including defects in branching morphogenesis in the lungs and salivary glands as well as epidermal hypoplasia and reduced hair follicles (Bebee et al. 2015). These findings suggest some functional redundancy between ESRP1 and ESRP2.

6 Concluding Remarks

Recent findings have revealed that ESRPs can either positively or negatively impact cancer progression. Thus, the plastic nature of their expression as well as the fine-tuning of their activities appears to be prerequisites for successful cancer progression. Several possible molecular mechanisms for regulating the expression and activity of ESRPs, have emerged.

First, ESRPs expression is regulated in complex circuits comprising transcriptional repressors and miRNAs. Notably, a double negative feedback circuit between ESRPs and δEF1/SIP1 to turn on/off the switch of ESRP expression in some types of cells can play a crucial role in the plastic expression of ESRPs (Fig. 1). Identification of internal or external cues that affect the balance in the circuit would help further our understanding of how this intricate regulatory system operates *in vivo*. The underlying mechanisms may not be uniform, as the circuit does not appear to operate in other cells.

Second, the activity of ESRPs can also be modulated through multiple mechanisms. Recently hnRNPM was reported to be a functional antagonist of ESRPs (Xu et al. 2014). In contrast to ESRPs, hnRNPM is a splicing regulatory protein that is highly expressed in mesenchymal cells but downregulated in epithelial cells. It drives splicing programs that oppose those promoted by ESRPs. There may be other functionally antagonistic splicing regulators that remain to be identified. Post-translational modifications of ESRPs appear to be important but only poorly understood. Thus far, only the K27-linked polyubiquitination of ESRP2 by Arkadia has been shown to enhance the splicing function of ESRP2 (Mizutani et al. 2015). Other covalent modifications of ESRP proteins could modulate their functions as well, which remain to be elucidated.

Given the dual functions of ESRPs in cancer progression, simple enhancement or inhibition of their activities would have unfavorable outcomes in patients. ESRPs themselves may therefore not be suitable for therapeutic molecular targets. The same argument can be applied to the use of ESRPs as prognostic markers, which has given controversial results thus far (Ueda et al. 2014; Mizutani et al. 2015; Preca et al. 2015; Yae et al. 2012). In addition to the plastic nature of ESRP1/2 expression during cancer progression (Ueda et al. 2014; Ishii et al. 2014), post-translational modification of ESRP2 is required at least under some conditions (Mizutani et al. 2015), indicating that expression of ESRP2 does not always correspond with its activity. The same may be true for ESRP1. Rather, proteins downstream of ESRPs could serve as prognostic markers (Mizutani et al. 2015) or molecular targets for therapeutic inhibition of cancer progression. Key molecular effectors downstream of ESRPs during cancer progression need to be further explored in the near future.

Conflict of Interest None declared

References

Bebee TW, Park JW, Sheridan KI, Warzecha CC, Cieply BW, Rohacek AM, Xing Y, Carstens RP (2015) The splicing regulators Esrp1 and Esrp2 direct an epithelial splicing program essential for mammalian development. eLife 4:e08954

Bhate A, Parker DJ, Bebee TW, Ahn J, Arif W, Rashan EH, Chorghade S, Chau A, Lee JH, Anakk S, Carstens RP, Xiao X, Kalsotra A (2015) ESRP2 controls an adult splicing programme in hepatocytes to support postnatal liver maturation. Nat Commun 6:8768

Braunschweig U, Gueroussov S, Plocik AM, Graveley BR, Blencowe BJ (2013) Dynamic integration of splicing within gene regulatory pathways. Cell 152:1252–1269

Brown RL, Reinke LM, Damerow MS, Perez D, Chodosh LA, Yang J, Cheng C (2011) CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression. J Clin Invest 121:1064–1074

Cieply B, Farris J, Denvir J, Ford HL, Frisch SM (2013) Epithelial-mesenchymal transition and tumor suppression are controlled by a reciprocal feedback loop between ZEB1 and Grainyhead-like-2. Cancer Res 73:6299–6309

Comijn J, Berx G, Vermassen P, Verschueren K, van Grunsven L, Bruyneel E, Mareel M, Huylebroeck D, van Roy F (2001) The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. Mol Cell 7:1267–1278

David CJ, Manley JL (2010) Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. Genes Dev 24:2343–2364

Dittmar KA, Jiang P, Park JW, Amirikian K, Wan J, Shen S, Xing Y, Carstens RP (2012) Genome-wide determination of a broad ESRP-regulated posttranscriptional network by high-throughput sequencing. Mol Cell Biol 32:1468–1482

Eger A, Aigner K, Sonderegger S, Dampier B, Oehler S, Schreiber M, Berx G, Cano A, Beug H, Foisner R (2005) DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. Oncogene 24:2375–2385

Eswarakumar VP, Lax I, Schlessinger J (2005) Cellular signaling by fibroblast growth factor receptors. Cytokine Growth Factor Rev 16:139–149

Fagoonee S, Bearzi C, Di Cunto F, Clohessy JG, Rizzi R, Reschke M, Tolosano E, Provero P, Pandolfi PP, Silengo L, Altruda F (2013) The RNA binding protein ESRP1 fine-tunes the expression of pluripotency-related factors in mouse embryonic stem cells. PLoS One 8:e72300

Fiegen D, Haeusler LC, Blumenstein L, Herbrand U, Dvorsky R, Vetter IR, Ahmadian MR (2004) Alternative splicing of Rac1 generates Rac1b, a self-activating GTPase. J Biol Chem 279:4743–4749

Horiguchi K, Sakamoto K, Koinuma D, Semba K, Inoue A, Inoue S, Fujii H, Yamaguchi A,

Miyazawa K, Miyazono K, Saitoh M (2012) TGF-β drives epithelial-mesenchymal transition through δEF1-mediated downregulation of ESRP. Oncogene 31:3190–3201

Horvath A, Pakala SB, Mudvari P, Reddy SD, Ohshiro K, Casimiro S, Pires R, Fuqua SA, Toi M, Costa L, Nair SS, Sukumar S, Kumar R (2013) Novel insights into breast cancer genetic variance through RNA sequencing. Sci Rep 3:2256

Hoshi Y, Endo K, Shirakihara T, Fukagawa A, Miyazawa K, Saitoh M (2016) The potential role of regulator of G-protein signaling 16 in cell motility mediated by δEF1 family proteins. FEBS Lett 590:270–278

Ishii H, Saitoh M, Sakamoto K, Kondo T, Katoh R, Tanaka S, Motizuki M, Masuyama K, Miyazawa K (2014) Epithelial splicing regulatory proteins 1 (ESRP1) and 2 (ESRP2) suppress cancer cell motility via different mechanisms. J Biol Chem 289:27386–27399

Ishimoto T, Nagano O, Yae T, Tamada M, Motohara T, Oshima H, Oshima M, Ikeda T, Asaba R, Yagi H, Masuko T, Shimizu T, Ishikawa T, Kai K, Takahashi E, Imamura Y, Baba Y, Ohmura M, Suematsu M, Baba H, Saya H (2011) CD44 variant regulates redox status in cancer cells by stabilizing the xCT subunit of system xc(-) and thereby promotes tumor growth. Cancer Cell 19:387–400

Jordan P, Brazåo R, Boavida MG, Gespach C, Chastre E (1999) Cloning of a novel human Rac1b splice variant with increased expression in colorectal tumors. Oncogene 18:6835–6839

Kaida D, Schneider-Poesch T, Yoshida M (2012) Splicing in oncogenesis and tumor suppression. Cancer Sci 103:1611–1616

Kalluri R, Weinberg RA (2009) The basics of epithelial-mesenchymal transition. J Clin Invest 119:1420–1428

Leontieva OV, Ionov Y (2009) RNA-binding motif protein 35A is a novel tumor suppressor for colorectal cancer. Cell Cycle 8:490–497

Lu H, Liu J, Liu S, Zeng J, Ding D, Carstens RP, Cong Y, Xu X, Guo W (2013) Exo70 isoform switching upon epithelial-mesenchymal transition mediates cancer cell invasion. Dev Cell 27:560–573

Mizutani A, Koinuma D, Seimiya H, Miyazono K (2015) The Arkadia-ESRP2 axis suppresses tumor progression: analyses in clear-cell renal cell carcinoma. Oncogene. doi:10.1038/onc2015.412

Ponta H, Sherman L, Herrlich PA (2003) CD44: from adhesion molecules to signalling regulators. Nat Rev Mol Cell Biol 4:33–45

Preca BT, Bajdak K, Mock K, Sundararajan V, Pfannstiel J, Maurer J, Wellner U, Hopt UT, Brummer T, Brabletz S, Brabletz T, Stemmler MP (2015) A self-enforcing CD44s/ZEB1 feedback loop maintains EMT and stemness properties in cancer cells. Int J Cancer 137:2566–2577

Reinke LM, Xu Y, Cheng C (2012) Snail represses the splicing regulator epithelial splicing regulatory protein 1 to promote epithelial-mesenchymal transition. J Biol Chem 287:36435–36442

Schnelzer A, Prechtel D, Knaus U, Dehne K, Gerhard M, Graeff H, Harbeck N, Schmitt M, Lengyel E (2000) Rac1 in human breast cancer: overexpression, mutation analysis, and characterization of a new isoform, Rac1b. Oncogene 19:3013–3020

Shirakihara T, Horiguchi K, Miyazawa K, Ehata S, Shibata T, Morita I, Miyazono K, Saitoh M (2011) TGF-β regulates isoform switching of FGF receptors and epithelial-mesenchymal transition. EMBO J 30:783–795

Ueda J, Matsuda Y, Yamahatsu K, Uchida E, Naito Z, Korc M, Ishiwata T (2014) Epithelial splicing regulatory protein 1 is a favorable prognostic factor in pancreatic cancer that attenuates pancreatic metastases. Oncogene 33:4485–4495

Warzecha CC, Carstens RP (2012) Complex changes in alternatice pre-mRNA splicing play a central role in the epithelial-to-mesenchymal transition (EMT). Semin Cancer Biol 22:417–427

Warzecha CC, Sato TK, Nabet B, Hogenesch JB, Carstens RP (2009a) ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing. Mol Cell 33:591–601

Warzecha CC, Shen S, Xing Y, Carstens RP (2009b) The epithelial splicing regulatory factors ESRP1 and ESRP2 positively and negatively regulate diverse types of alternative splicing events. RNA Biol 6:546–562

Warzecha CC, Jiang P, Amirikian K, Ditmmar KA, Lu H, Shen S, Guo W, Xing Y, Carstens RP (2010) An ESRP-regualted splicing programme is abrogated during the epithelial-mesenchymal transition. EMBO J 29:3286–3300

Werner S, Frey S, Riethdorf S, Schulze C, Alawi M, Kling L, Vafaizadeh V, Sauter G, Terracciano L, Schumacher U, Pantel K, Assmann V (2013) Dual roles of the transcription factor grainyhead-like 2 (GRHL2) in breast cancer. J Biol Chem 288:22993–23008

Xiang X, Deng Z, Zhuang X, Ju S, Mu J, Jiang H, Zhang L, Yan J, Miller D, Zhang HG (2012) Grhl2 determines the epithelial phenotype of breast cancers and promotes tumor progression. PLoS One 7:e50781

Xu Y, Gao XD, Lee JH, Huang H, Tan H, Ahn J, Reinke LM, Peter ME, Feng Y, Gius D, Siziopikou KP, Peng J, Xiao X, Cheng C (2014) Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing. Genes Dev 28:1191–1203

Yae T, Tsuchihashi K, Ishimoto T, Motohara T, Yoshikawa M, Yoshida GJ, Wada T, Masuko T, Mogushi K, Tanaka H, Osawa T, Kanki Y, Minami T, Aburatani H, Ohmura M, Kubo A, Suematsu M, Takahashi K, Saya H, Nagano O (2012) Alternative splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell. Nat Commun 3:883

Controlling Autolysis During Flagella Insertion in Gram-Negative Bacteria

Francesca A. Herlihey and Anthony J. Clarke

Abstract

The flagellum is an important macromolecular machine for many pathogenic bacteria. It is a hetero-oligomeric structure comprised of three major sub-structures: basal body, hook and thin helical filament. An important step during flagellum assembly is the localized and controlled degradation of the peptidoglycan sacculus to allow for the insertion of the rod as well as to facilitate anchoring for proper motor function. The peptidoglycan lysis events require specialized lytic enzymes, β-*N*-acetylglucosaminidases and lytic transglycosylases, which differ in flagellated proteobacteria. Due to their autolytic activity, these enzymes need to be controlled in order to prevent cellular lysis. This review summarizes are current understanding of the peptidoglycan lysis events required for flagellum assembly and motility with a main focus on Gram-negative bacteria.

Keywords

Flagella • Peptidoglycan • Lytic transglycosylases • β-*N*-acetylglucosaminidases

Abbreviations

PG	peptidoglycan
MurNAc	*N*-acetylmuramic acid
GlcNAc	*N*-acetylglucosamine
PBP	penicillin-binding proteins
LT	lytic transglycosylase
T3SS	type three secretion system
Mot	motor
OmpA	outer membrane protein A
PGB	peptidoglycan binding
Mlt	membrane-bound lytic transglycosylases
Slt	soluble lytic transglycosylases
GH	glycoside hydrolase

F.A. Herlihey and A.J. Clarke (✉)
Department of Molecular and Cellular Biology, University of Guelph, Guelph, ON N1G2W1, Canada
e-mail: aclarke@uoguelph.ca

StFlgJ	Salmonella Typhimurium FlgJ
PDB	Protein Data Bank
SAC	substrate-assisted catalysis
RlpA	rare lipoprotein A
1,6-anhydroMurNAc	1,6-anhydromuramic acid
NAG thiazoline	N-acetylglucosamine thiazoline.

1 Introduction

Peptidoglycan (PG) is an essential and unique structural element of the cell wall in nearly all bacteria. This heteropolymer is made up of glycan strands and peptide chains forming a continuous, mesh-like layer which surrounds the bacterial cell to confer strength, support and shape, as well as resistance to internal turgor pressure. The glycan strands are composed of repeating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) residues linked by β-1,4 glycosidic bonds. These strands are then cross-linked together through linkages between stem peptides attached to the lactyl groups of MurNAc residues. This combination of both glycosidic and peptide linkages between the PG components generates a single macromolecule, or sacculus, that completely surrounds the cell to maintain the integrity of the cytoplasmic membrane. Once formed, the PG sacculus is not a static structure as it requires constant remodeling and reinforcement to permit cellular growth, division and insertion of cellular machinery, such as the flagellum. The continuous biosynthetic events that occur external to the cytoplasmic membrane involve the coordinated action of PG-synthesizing enzymes, the penicillin-binding proteins (PBPs), and PG-degrading enzymes, such as β-N-acetylglucosaminidases and lytic transglycosylases (LTs) (reviewed in (Vollmer and Bertsche 2008)). LTs (EC 4.2.2.n1/n2) have the same substrate specificity as the muramidases (lysozymes; EC 3.2.1.17), viz. lysing the glycosidic bond between MurNAc and GlcNAc, but with a different reaction mechanism that does not

involve water. Thus, instead of producing a reducing hydrolytic product, the LTs lyse PG with the concomitant formation of an intramolecular 1,6-anhydromuramoyl reaction product (Fig. 1) (Höltje et al. 1975). The β-N-acetylglucosaminidases (EC 3.2.1.14; chitinase) hydrolyze the other glycosidic bond of PG, that between GlcNAc and MurNAc, generating a reducing GlcNAc product (Fig. 1) (Vollmer et al. 2008). This review presents our current understanding of the PG lysis events for flagellar insertion catalyzed by these PG lytic enzymes, with a focus on Gram-negative bacteria. Only proteins of relevance will be highlighted within this review.

2 Bacterial Flagellum Biosynthesis

Many bacteria are motile through the use of the flagellum; a complex nanomachine whose assembly is dependent on multiple cooperating components. It extends from the cytoplasm to the cellular exterior requiring approximately 25 different proteins which are variable in stoichiometry and are arranged into three major substructures: basal body, hook and thin helical filament (Fig. 2). The basal body consists of the rotary motor, a flagellum-specific type III secretion system (T3SS), and at least four ring-like structures which are all connected by a filamentous rod (reviewed in (Macnab 2003)).

The assembly of the bacterial flagellum has been extensively reviewed (eg. Minamino and Imada 2015) and so the following presents only a brief overview. Assembly is initiated with the formation of the MS-ring within the cytoplasmic membrane, followed by attachment of the switch complex (C-ring) at the cytoplasmic face of the MS-ring. Once completed, the intrinsic membrane motor (Mot) proteins MotA and MotB assemble to form the stator complex which creates the pathway for proton influx across the cytoplasmic protein. The MotA/MotB complex is believed to be anchored to the PG layer by a transient interaction through MotB's putative PG-binding (PGB) motif located within the
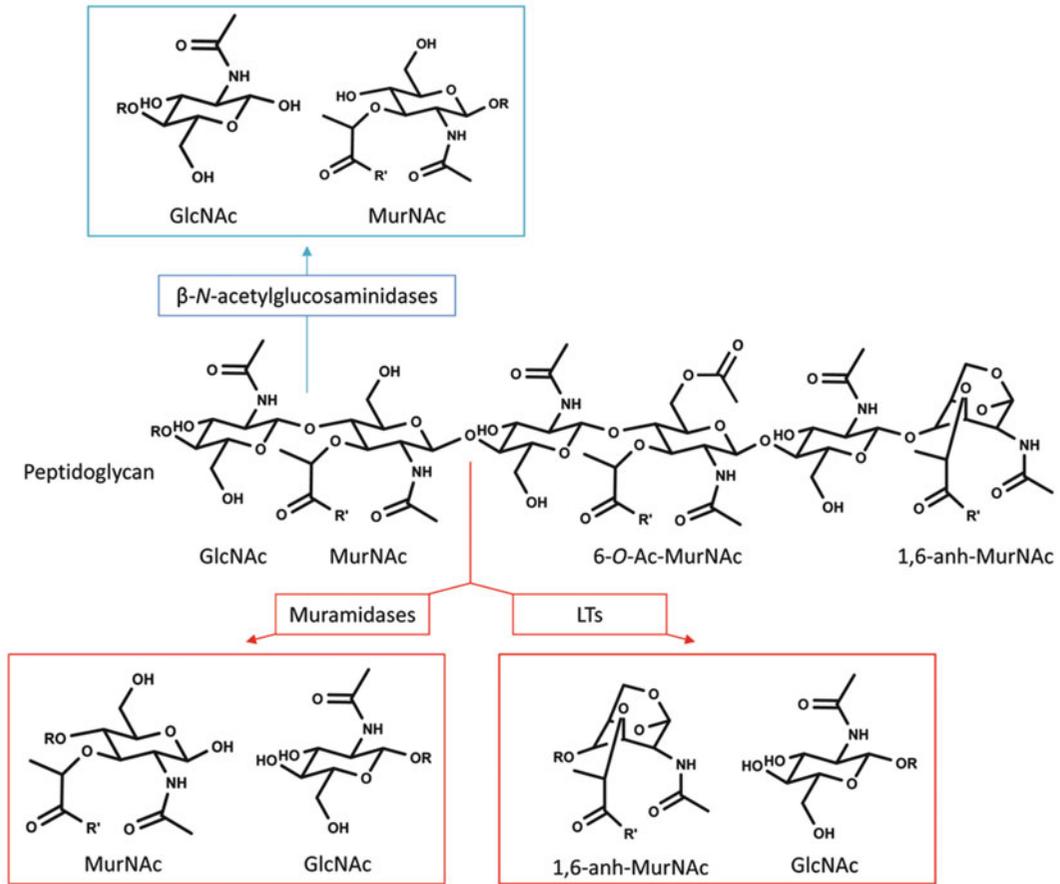
Fig. 1 Structure of PG and reactions catalyzed by glycolytic autolysins. R and R' denote MurNAc and stem peptides, respectively (The figure is adapted from Ref. (Herlihey et al. 2014))

C-terminal periplasmic domain (De Mot and Vanderleyden 1994). Recently determined crystal structures of MotB from *Helicobacter pylori* (Roujeinikova 2008) and *Salmonella* Typhimurium (Kojima et al. 2008a) revealed that these proteins share a common outer membrane protein A-like (OmpA) domain known to bind to the PG layer. Furthermore, it has been shown that the PGB region of Pal, a PG-associated lipoprotein, is interchangeable with the PGB region of MotB (Hizukuri et al. 2009). Proper anchoring to the PG sacculus requires a relatively large conformational change within the periplasmic domain of MotB (Reboul et al. 2011; Hoskin et al. 2006; Kojima and Blair 2001). Once properly anchored into the sacculus, the MotA/MotB complex switches from an inactive ('plug' closed) state to an active

('unplugged' open) state to open the proton pathway upon assembly of the complex to the motor (Reboul et al. 2011; Hoskin et al. 2006; Kojima and Blair 2001).

The flagellum specific T3SS assembles at the base of the flagellar basal body within the MS-ring and consists of six integral membrane proteins and three cytoplasmic proteins (Suzuki et al. 1998; Minamino and Macnab 1999). The apparatus facilitates the proton motive force-dependent export of most components of the flagellum (Koushik et al. 2008; Minaminoa and Namba 2008) with the exception of the P-ring protein (FlgI) and L-ring lipoprotein; these two ring proteins are exported *via* the Sec-dependent pathway (Homma et al. 1987). The filamentous rod extends from the MS-ring through the P- and L-rings and it is divided into a proximal

Fig. 2 Organization of the rod and motor proteins of the bacterial flagellum within the Gram-negative bacterial envelope. *IM* inner membrane, *OM* outer membrane, *PG* peptidoglycan
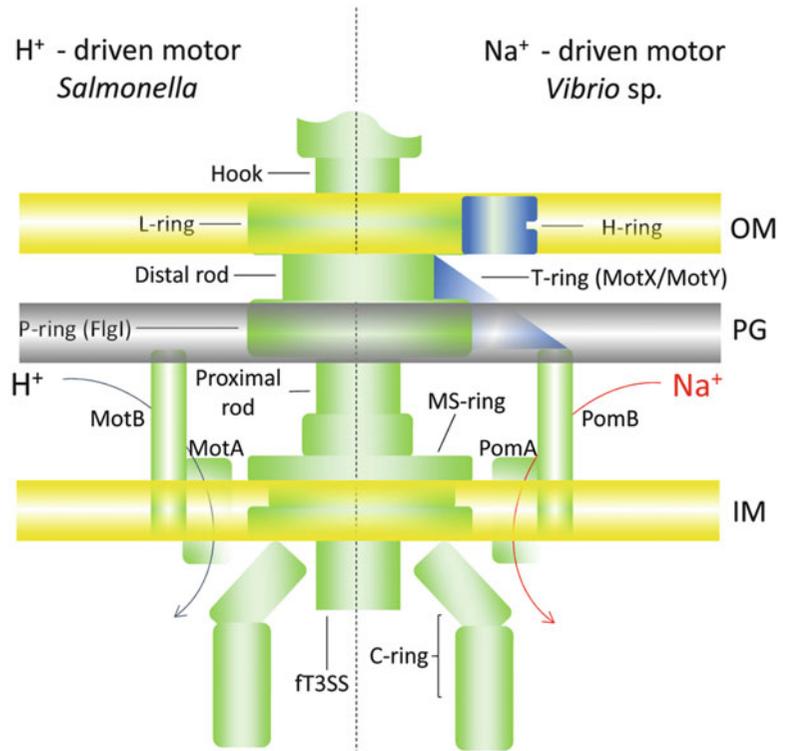
(involving FliE, FlgB, FlgF and FlgC) and distal (FlgG) rod which is progressively assembled (Osorio-Valeriano et al. 2016). Assembly of the P- and L-rings are reliant on the construction of the preceding flagellar rod structure (Kubori et al. 1992). The P-ring is believed to be fixed to the PG layer, whereas the L-ring resides in the outer membrane. Together, these form a stiff cylindrical structure which is understood to act as a bushing for rod rotation (Aika et al. 1990). Once the rod and rings are completed, the hook and filament are assembled, the latter involving the non-covalent polymerization of flagellin proteins (Macnab 2003).

The basic structure of the basal body is conserved amongst bacterial species, however there exist variations depending on the ions used for motive force. In some bacteria, such as *Vibrio sp.*, sodium ions are used instead of protons which requires the homologous motor proteins PomA and PomB, as well as two additional rings, the T- and H-ring (Minamino and Imada 2015) (Fig. 2). Similar to MotB, the

C-terminal periplasmic region of PomB contains a putative PGB motif (Asai et al. 1997) and it is thought to anchor the stator to the PG layer after undergoing a large conformational change (Zhu et al. 2014). Beneath the P-ring is the T-ring which is composed of MotX and MotY. These two proteins are essential for incorporating the stator unit into the motor and then stabilizing it (Terashima et al. 2006). The crystal structure of MotY has been solved and it has high similarity to outer membrane protein A (Kojima et al. 2008b). In addition, it contains a disordered PGB motif which has been proposed to prevent MotY from binding tightly to the PG layer before the MotX/MotY complex encounters the basal body. Interaction of the MotX/MotY complex with the basal body may cause the disordered chains to fold into a functional PGB pocket (Kojima et al. 2008b). Lastly, the H-ring surrounds the L- and P-rings and it too is also required for proper rod assembly of the stator units around the rotor (Terashima et al. 2010).

3 Peptidoglycan Degrading Enzymes

3.1 Role in Flagella Assembly

Extensive modifications to the PG sacculus need to be made during basal body formation to accommodate the insertion of the apparatus and then to stabilize the function of this system by acting as an assembly scaffold. An important step in basal body formation is the penetration of the rod through the PG layer. The diameter of the rod is approximately 8–14 nm which is larger than the average pore size of 2–4 nm within the PG sacculus (Suzuki et al. 1998; Zhao et al. 2013; Demchick and Koch 1996; Vázquez-Laslop et al. 2001; Gumbart et al. 2014). Hence, its insertion requires the localized and controlled lysis of a limited number of PG strands. In species of the β- and γ-proteobacteria, this degradation is performed by FlgJ, a bifunctional protein with an N-terminal domain responsible for proper rod assembly (Hirano et al. 2001) and a C-terminal domain possessing β-N-acetylglucosaminidase activity (Herlihey et al. 2014). However, a proteomics study suggested that FlgJ homologues in a number of flagellated α-proteobacteria appear to lack this C-terminal PG lytic domain (Nambu et al. 2006; González-Pedrajo et al. 2002). Using Rhodobacter sphaeroides as a model bacterium, this lack of FlgJ lytic activity was shown to be compensated for by a specialized lytic enzyme (de la Mora et al. 2007) which was subsequently demonstrated experimentally to be an endo-acting LT, SltF. Homologs of SltF appear to be widespread in flagellated α-proteobacteria that lack a bimodular FlgJ (Herlihey et al. 2016). An exception amongst the α-proteobacteria is Caulobacter crescentus which lacks an SltF homolog. Instead, PleA, an uncharacterized PG degrading enzyme, is proposed to provide the necessary PG lytic activity for flagella and pili assembly in this bacterium (Viollier and Shapiro 2003).

A crucial feature of the flagellum for full motility is its anchoring to the sacculus through MotB (PomB). For this to occur, it has been postulated that MotB requires the localized remodelling of PG. This is accomplished in Gram-negative bacteria by the LTs, such as membrane-bound LT D (MltD) from H. pylori, as well as soluble LT (Slt) and MltC from S. Typhimurium (Roure et al. 2012). Thus, the LTs may serve the dual function of boring the hole within the sacculus and then restructuring the PG around its circumference. Moreover, the specific 1,6-anhydromuramoyl reaction products of this activity may provide binding sites for MotB to tether the flagellum to the sacculus. Indeed, the lack of a functional MltD in H. pylori led to a loss of motility even though the flagellum was fully assembled (Roure et al. 2012). With Gram-positive bacteria, this remodelling of their multi-layered sacculus is catalysed by β-N-acetylglucosaminidases, such as Auto from Listeria monocytogenes (Roure et al. 2012) or LytD in conjunction with LytC, an N-acetylmuramoyl-L-Ala amidase, in Bacillus subtilits (Chen et al. 2009). With these cells, MotB is expected to be tethered to the reducing GlcNAc and/or non-reducing MurNAc reaction products but the nature of these interactions are not known. The importance of this lytic activity was demonstrated in knockout studies of the genes encoding their respective lytic enzymes which, as with H. pylori, led to immotile cells possessing fully assembled flagella (Roure et al. 2012; Chen et al. 2009). It should be noted that both OmpA and Pal have been demonstrated experimentally to bind to the stem peptide of PG through a non-covalent interaction with meso-diaminopimelic acid and it has been proposed that MotB also binds this moiety in the stem peptide along with the glycan chain (Roujeinikova 2008; Reboul et al. 2011; Parsons et al. 2006; Park et al. 2012). Currently, it is unknown what muropeptides FlgI recognize and bind to.

3.2 β-N-Acetylglucosaminidases

3.2.1 Families

The Carbohydrate-Active enZyme (CAZy) database categorizes β-N-acetylglucosaminidases into seven different glycoside hydrolase (GH) families, GH3, GH18, GH20, GH56,

GH73, GH84 and GH85, based on their secondary structure predictions (Lombard et al. 2014). Members classified into GH73 (Pfam PF01832) are involved in hydrolyzing PG and include the β-N-acetylglucosaminidases involved in flagellum assembly and motility. The family has been organized into five phylogenetic clusters harbouring characteristic sequence motifs among three different phyla. All sequences within clusters 1 and 5 belong to proteobacteria and contain enzymes that do not carry additional domains within their catalytic modules. The sequences found within clusters 2 and 4 belong to firmicutes and contain enzymes which are highly modular while all sequences within cluster 3 belong to bacteroidetes (Lipski et al. 2015). Located within cluster 1 is FlgJ. Recently, *S.* Typhimurium FlgJ (*St*FlgJ) was identified to have nine consensus motifs in which the majority of these (five) are associated with the C-terminal β-N-acetylglucosaminidase domain (Herlihey et al. 2014). The two putative catalytic residues, Glu184 and Glu223, are found within motif IV and motif VI, respectively, and motif IX contains a Tyr-Ala-Thr-Asp sequence which is conserved in all clusters except cluster 5 (Herlihey et al. 2014; Lipski et al. 2015).
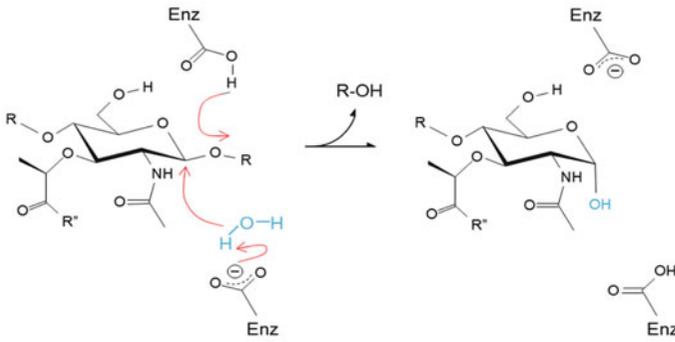
3.2.2   Mechanism of Action

Most GH enzymes catalyze either single- or double- displacement reactions involving the direct participation of two catalytic residues (Glu and/or Asp) which are appropriately spaced across the active site (Fig. 3). In the single-displacement mechanism, a catalytic residue acts as a general acid to protonate the glycosidic oxygen. The second residue is positioned across the active-site cleft by an average distance of 10 Å and acts as a general base to abstract a proton from water generating a hydroxide ion. This ion attacks the anomeric C-1 carbon of the GlcNAc residue, resulting in bond cleavage and a product with an inverted anomeric stereochemistry. As its name implies, the double-displacement reaction typically involves two distinct reactions with a covalent intermediate separating them. In the first reaction, and as with the single displacement mechanism, a catalytic acid protonates the glycosidic oxygen to be cleaved while the second acidic catalytic residue serves as a nucleophile to attack the anomeric center of the oxocarbenium ion transition state leading to departure of the non-reducing end product and covalent attachment of the remaining substrate. In the second half of the reaction, the catalytic acid now functions as a base to abstract a proton from a water and the resulting hydroxide ion attacks the covalent adduct leading to its hydrolysis from the enzyme. Due to the positioning and direction of the two nucleophilic attacks, first by the catalytic nucleophile and then the hydroxide, the anomeric configuration of the resulting GlcNAc product is retained.

With the GH73 β-N-acetylglucosaminidases, the issue of its catalytic mechanism is not so clear. Studies involving the site-specific replacement of conserved amino acid residues agreed that Glu184 (*S.* Typhimurium FlgJ numbering) was essential for catalytic activity (Inagaki et al. 2009; Maruyama et al. 2010; Yokoi et al. 2008). However, conflicting results failed to identify unambiguously a second catalytic residue which has led to some controversy. Some insight on this situation was gained following the determination of the crystal structures of several enzymes. The structures of five GH73 enzymes are now known, those of *L. monocytogenes* Auto (PDB 3F17), *Streptococcus pneumoniae* LytB (PDB 4Q2W), *Thermotoga maritima* TM0663 (PDB 4QDN), *Sphingomonas* sp. FlgJ (*Sp*FlgJ) (PDB 2ZYC), and *St*FlgJ (PDB 5DN4). Each possesses an α/β-hydrolase-fold with a conserved active site and a deep cleft that accommodates their PG substrates. The homologous Glu184 residues were observed to be appropriately positioned within the active-site cleft to function as the acid catalyst. However, identification of a second conserved active-site acidic residue could not be made for each of these enzymes. Such a second residue was observed at the putative catalytic centres of TM0663, Auto, *Sp*FlgJ and *St*FlgJ but quite distant from their respective catalytic acids (Lipski et al. 2015; Hashimoto et al. 2009; Bublitz et al. 2009; Zabola et al. 2016). With Auto, *Sp*FlgJ and *St*FlgJ, the
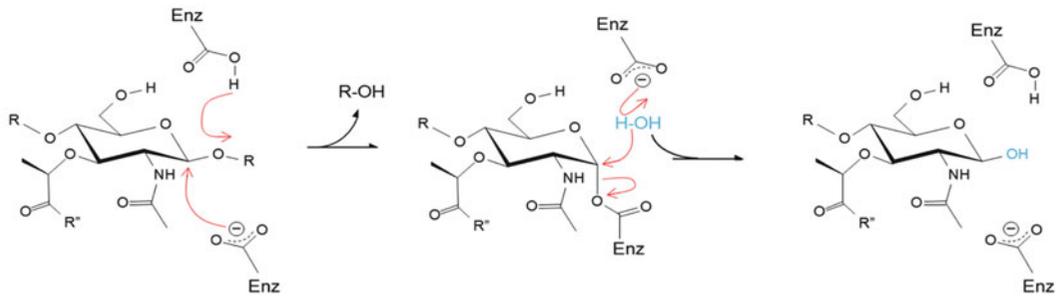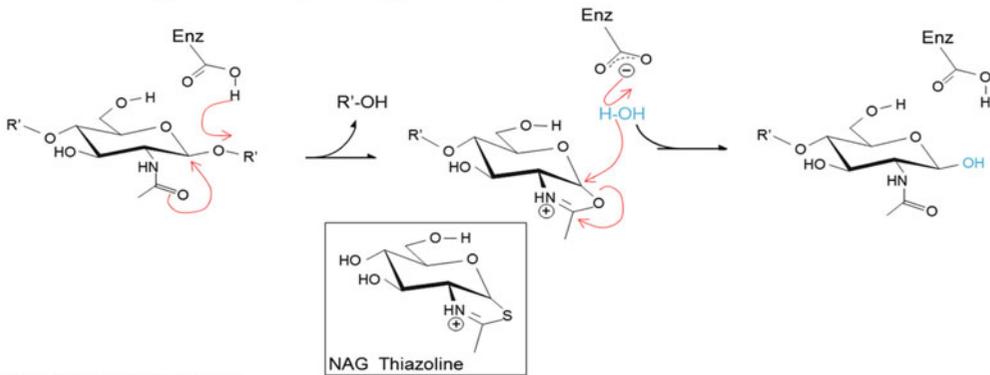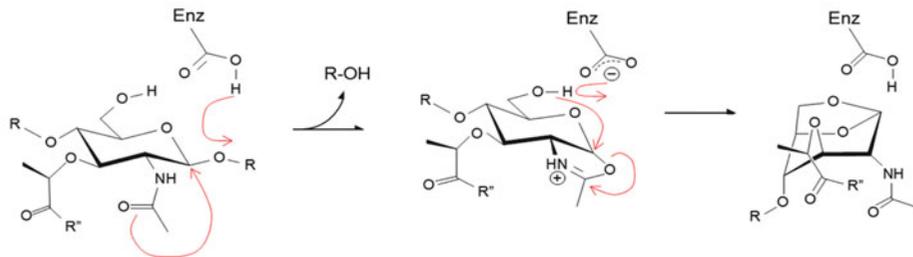
Fig. 3 Mechanism of action of glycolytic enzymes active on PG. R, R', and R'' denote GlcNAc, MurNAc, and stem peptides, respectively. *Inset*: Structure of the β-*N*-acetylglucosaminidase inhibitor *N*-acetylglucosamine (NAG) thiazoline

putative acid catalysts are positioned on an α-helix while the second proposed catalytic glutamyl residues are located on a flexible β-hairpin. The separation of these residues is 13 Å for Auto (Bublitz et al. 2009) and *Sp*FlgJ (Maruyama et al. 2010), and 21.6 Å for *St*FlgJ (Zabola et al. 2016) which represents an outer limit for the distance between catalytic residues in an inverting enzyme (McCarter and Withers 1994). However, it has been postulated that this flexible β-hairpin undergoes a conformational change upon binding substrate in which the active site opens to capture the substrate and then closes to correctly position the active-site residues for catalysis (Maruyama et al. 2010; Bublitz et al. 2009; Zabola et al. 2016). Based on the eventual distances between the respective catalytic partners, these β-*N*-acetylglucosaminidases are still expected to catalyze an inverting mechanism. It remains to be established whether or not the β-hairpin does close to permit an inverting mechanistic pathway, if indeed this occurs recognizing that experimental evidence of this stereochemical outcome is lacking. Obtaining such evidence is central to this debate because an equally plausible outcome of the conformational shift is to position the second catalytic glutamate to assist the deprotonation of the amide of an *N*-acetyl group in an anchimeric mode of action resulting in the retention of anomeric configuration in the lytic product. Indeed, this alternative mechanism of action, also known as substrate-assisted catalysis (SAC), has been postulated for several other GH73 β-*N*-acetylglucosaminidases. With the apparent absence of a second catalytic residue in the determined crystal structure of LytB, SAC was invoked for glycoside lysis (Bai et al. 2014; Rico-Lastres et al. 2015). Based on molecular and biochemical studies combined with theoretical considerations, FlgJ (Herlihey et al. 2014), AltWM (Yokoi et al. 2008), and AcmA (Inagaki et al. 2009) were also proposed to catalyze a SAC mechanism of action. In this mechanism, the carbonyl oxygen of the GlcNAc N-acetyl group substitutes for the nucleophilic catalytic residue in the enzyme to generate an

intramolecular oxazolinium ion intermediate that stabilizes the oxocarbenium ion transition state formed by acid catalysis during glycolytic cleavage (Fig. 3). As with the second phase of the typical double-displacement reaction, a hydroxide generated from water hydrolyzes the oxazolinium ion to generate the reducing end GlcNAc product with retention of configuration. Support for this mechanism is provided by investigations with the closely related GH18 and GH20 glycoside hydrolases, many of which are also β-*N*-acetylglucosaminidases. These enzymes also do not have the typical second catalytic residue and they are inhibited by mechanism-based thiazoline inhibitors (Fig. 3 inset) (Knapp et al. 1996; Macdonald et al. 2010). With the GH73 enzymes, assistance in rendering the carbonyl oxygen nucleophilic has been proposed to be provided by a highly conserved Tyr residue positioned appropriately to deprotonate the *N*-acetamido nitrogen. Indeed, replacement of this Tyr residue in FlgJ, AcmA and AltWN resulted in reduced activity similar to that associated with the replacement of the general acid Glu (Inagaki et al. 2009; Maruyama et al. 2010; Yokoi et al. 2008).

3.3 Lytic Transglycosylases

3.3.1 Families

A classification scheme was developed to organize LTs into four distinct families based on their sequence similarities and identified consensus motifs. The archetypes for the respective families were represented by seven known LTs (Slt70, MltA-MltF) from *E. coli*, one LT (MltB) from *Pseudomonas aeruginosa* and the endolysins from λ-bacteriophage (Blackburn and Clarke 2001; Scheurwater et al. 2008). As the discovery of new LTs involved in specific tasks, such as *R. sphaeroides* SltF (Herlihey et al. 2016), *E. coli* MltG (Yunck et al. 2016) and *P. aeruginosa* rare lipoprotein A (RlpA) (Jorgenson et al. 2014), is becoming more frequent, an updated classification scheme has been generated here to include an additional subfamily 1F in family 1, as well as families 5 and 6 (Fig. 4).

Family   **Consensus Motifs**

	I	II	III	IV
1A	RQES·F·P·A·S·	·GA·GLMQ·MP·TA· (33-44)	A·AAYNAGPGRV (3-19)	E·IPF·ETR·YV··V
1B	·TES·FNPYA·S·	··A·GLMQVVP··AG (37-53)	·I·AYNGG·G·V (0-28)	T··PSA·SRRYL·KV
1C	AIESGGNPNAVSK	SNAIGLMQIKASTSG (37-53)	LVVSYANGAGAL (28)	·NHPA·QAPRYIWKL
1D	IVESA·F·F·A·S·	··A·GLWQFMP·TGR (38-42)	·AAYNEG·G·V (12-20)	L·LP··ET··YVPK·
1E	YQESHW···A·SPT	G·RG·MMLTK·TA· (39-42)	ALAAYN·G··H· (28-42)	GYARG·EA··YVENI

	I	II	III	IV	V
1F i	R·ESGR····G	AWPWT·NA·G·G·YF (20)	IDVGCMQINL··H (9)	AFDP··NA·YAA·FL (12)	AAAAYHS·TP··G·Y
ii	LTETGR····G	PWPWTVNMEG·G·WF (20)	FDVGCFQINYRWH (9)	MFDP··NA·YAARFL (12)	AAGAYHSRT···A·RY

	I	II	III	IV	V
2	DNF·M·VQGSGYID·GDG (6-11)	YAGKNG·PY··IGK·LIDK (6-7)	SMQAIR·W (3-17)	L·QNPS·VFFK (5-14)	PVKGAAAGVYPL····SVA·DR

(23-31) RLMVALDTGGAIKG····D·Y·G·G··AG··AG   VI

	I	II	III	IV
3	YGVP·IIV·IIGVET··G (9-10)	DALATL·F·Y (30-36)	GSYAGAMG··QFMPSSY (2)	YAVDFDGDGHI·LW··P·DA·GSVANYFK·HGW

	I	II	III	IV
4	A······SEG (6-12)	YDVIVGGELFTDYSDHPRKLV (11-15)	ST··A·GRYQ·L···W (45-63)	IWASLPGAGYGQ·EH

	I	II	III	IV
5	FKAGTYR·T (10)	LL·SGKEAQF··RLVEG·RL (36-38)	EG·F·PDT··YTA·TTD·ALLKRA (27)	T·MASIIEKETAV

(2) ER··VASVFINRLR·GMRLQTDVIYGMG (2)   VI   Y·G·L·R·DLETPT·YNTY·I·GLPP (9)   VII   SL·AAAHPAKT·Y·YFVADGKGGH·F·LASHN   VIII

	I	II	III	IV
6	TA·GE·FD (4)	TAAH·TLPLPSYV·VTN··NG (4)	VRVNDRGPF (3)	RVIDLS·AAA··L

Fig. 4 Family organization of the LTs. Residues in plain and boldface type of the consensus motifs (numbered with *roman numerals*) are present in greater than 80 % and invariant amongst the sequences of the individual families, whereas those in *red type* and labelled with the *asterisks* identify the putative catalytic acid/base residues. The *numbers in parentheses* denote the number of residues between the motifs

The following presents only a brief overview of the families originally identified by (Blackburn and Clarke 2001) (Blackburn and Clarke 2001). Initially, family 1 was a superfamily consisting of five subfamilies, with each sharing limited sequence similarity to goose-type lysozyme. Included within this family are hypothetical and known LTs with sequence identities to *E. coli* Slt70 (1A), MltC (1B), MltE (1C), MltD (1D) and MltF (1E). Recently, SltF encoded in flagellated α-proteobacteria, was identified to comprise a new subfamily, 1F. Members of this sub-family possess four consensus motifs and can be further divided into two subsets distinguished primarily by the presence of either a Ser or Thr residue immediately following the putative catalytic Glu in motif I (Herlihey et al. 2016). Currently, all of the known LTs involved in flagella assembly and insertion are found within this family. Family 2 is comprised of hypothetical and known enzymes which share identity to the *E. coli* enzyme MltA. Unlike the other families, these enzymes contain two conserved aspartates, both of which are important for catalysis although only one acts as the general acid/base (further discussed below). The third family of hypothetical and known LTs share similarities with *E. coli*

and *P. aeruginosa* MltB. In contrast to family 1 LTs, enzymes classified into this family contain a conserved catalytic Glu which is followed by an invariant Thr in motif I. Family 4 consists mainly of proteins of bacteriophage origin.

Two additional families, family 5 and 6, have been added to this classification scheme using *E. coli* MltG and *P. aeruginosa* RlpA, respectively, as the archetypes. Hypothetical and known enzymes which are classified into family 5 are highly homologous and contain eight consensus motifs (Fig. 4). Like members of family 3 enzymes, these LTs possess a conserved putative catalytic Glu followed by an invariant Thr but positioned further into the proteins, in motif IV. Family 6 enzymes have four consensus motifs with a conserved putative catalytic Asp found in motif III. With both families, further experimental evidence is required to confirm the role, if any, in catalysis and/or PG binding that the identified conserved residues in these motifs may have.

3.3.2　Mechanism of Action

Crystal structures have been solved for eleven LTs representing families: 1A (*E. coli* Slt70, PDB 1QSA), 1B (*E. coli* MltC, PDB 4C5F), 1C (*E. coli* MltE, PDB 2Y8P), 1E (*P. aeruginosa* MltF, PDB 4P11), 2 (*E. coli* MltA, PDB 2GAE; *Neisseria gonorrhoeae* MltA, PDB 2G6G), 3 (*E. coli* Slt35, PDB 1QUS; *P. aeruginosa* SltB1, PDB 4ANR; *P. aeruginosa* SltB3, PDB 5ANZ), 4 (bacteriophage lamba LT, PDB 1D9U), and 5 (*E. coli* MltG, PDB 2R1F). Despite limited sequence similarities between the family 1, 3, 4 and 5 LTs, they all share a common goose-type lysozyme fold in which the catalytic active site cleft is mainly α-helical (Thunnissen et al. 1995; Artola-Recolons et al. 2014; Artola-Recolons et al. 2011; van Asselt et al. 1999a; van Asselt et al. 2000; Nikolaidis et al. 2012; Lee et al. 2016; Leung et al. 2001). In contrast, both family 2 and 6 LTs are structured as a double-ψ-β-barrel which is reminiscent of an endoglucanase V fold (Jorgenson et al. 2014; van Straaten et al. 2005; Powell et al. 2006).

The majority of the LTs studied to date have the general ability to perform exo-lytic reactions which is accounted for by examination of the sub-site structure of the binding clefts of respective enzymes. Generally, these exo-lytic LTs, such as *E. coli* MltB, have four binding sub-sites (labeled −2, −1, +1, +2) that accommodate a pair of GlcNAc-MurNAc(peptide) residues where lysis occurs to release GlcNAc-1,6-anhydroMurNAc products from the "reducing" end of PG strands (van Asselt et al. 1999a; van Asselt et al. 2000). On the other hand, *E. coli* MltE, and likely MltF, are endo-lytic, cleaving within lengths of PG chains (Lee et al. 2013; Scheurwater and Clarke 2008). To accomplish this, these LTs have an open ended binding cleft that can accommodate at least eight saccharides (Artola-Recolons et al. 2011). Then there are LTs, such as *E. coli* MltC, that have a subsite architecture that accommodates up to nine aminosugars, and accordingly catalyze both endo- and exo-lytic activity (Artola-Recolons et al. 2014; Lee et al. 2013). In addition to the distinction, or not, of endo- and exo-activity, there is gathering evidence to suggest that family members share distinct substrate specificities. For example, family 6 enzymes appear to be active only on "naked" glycan strands that lack stem peptides (Jorgenson et al. 2014). Regardless of the discrete substrate specificity, cleavage of the PG chain occurs at the catalytic center of the LT that is positioned between the MurNAc and GlcNAc occupying the −1 and +1 subsites (Davies et al. 1997).

With the production of a 1,6-anhydromuramoyl product from the lysis of the β-1,4 glycosidic linkage between MurNAc and GlcNAc residues, the LTs in effect catalyze a retaining mechanism of action. However, only a single acidic residue is observed to be positioned appropriately at the putative catalytic centers in the known structures of the respective enzymes. Consequently, the LTs have been postulated to catalyze a SAC reaction analogous to that known for some β-1,4-*N*-acetylglucosaminidases (Terwisscha van Scheltinga et al. 1995; Tews et al. 1996; Drouillard et al. 1997; van Asselt et al. 1999b; Reid et al. 2004; Reid et al. 2007). As described above, the single catalytic Glu or Asp would act as a general acid to protonate the glycosidic oxygen of the linkage to be cleaved and an oxazolinium intermediate is

proposed to stabilize the putative oxocarbenium ion transition state. However, in this case, the oxazolinium would involve the N-acetyl group of MurNAc at subsite −1 and then the deprotonated catalytic residue acts as a general base to abstract a proton from the C-6 hydroxyl of this MurNAc residue. This promotes an intramolecular nucleophilic attack at the anomeric C-1 carbon of MurNAc collapsing the oxazolinium intermediate with the formation of the 1,6-anhydroMurNAc reaction product (Fig. 3). With the lack of defined substrates, it has not been possible to kinetically test this hypothesis regarding mechanism. However, preliminary evidence in support has been reported in a study involving the mechanism-based inhibitor of β-1,4-N-acetylglucosaminidases, NAG-thiazoline (Fig. 3, inset). Thus, this analog of the putative oxazolinium intermediate serves as an inhibitor of MltB from *P. aeruginosa*, albeit weakly, whereas neither inhibition by nor binding to its parent compound GlcNAc is observed (Reid et al. 2004).

Recently, a secondary minor muramidase-like activity has been reported for Slt70, MltC, MltD and MltE from *E. coli* (Artola-Recolons et al. 2014; Lee et al. 2013), SltF from *R. sphaeroides* (Herlihey et al. 2016) and CwlQ from *B. subtilis* (Sudiarta et al. 2010). It is postulated that these hydrolytic products arise as a water molecule gains access to the active center and it, instead of the C-6 hydroxyl of MurNAc, is deprotonated by the catalytic base and subsequently attacks the transient oxocarbenium ion in the second half of this double-displacement mechanism (Artola-Recolons et al. 2014; Lee et al. 2013).

4 Modulation of Flagella-Specific Autolysins

4.1 Control of Autolysins at the Enzyme Level

The LTs and β-N-acetylglucosaminidases represent major classes of autolytic enzymes that when left uncontrolled can cause complete cellular lysis. Given this, it is crucial for the cell to control these enzymes both temporally and spatially. Whereas little is understood about the β-N-acetylglucosaminidases, the activity of LTs is known to be controlled by three mechanisms at the enzyme level, their localization to membranes and associated protein complexes, modification of their PG substrate, and by the presence of proteinaceous inhibitors (reviewed in (Scheurwater et al. 2008; Moynihan and Clarke 2011)). There is growing evidence, particularly with the LTs involved in general PG biosynthesis, that these potentially autolytic enzymes complex with the PBPs and other proteins of the biosynthetic machinery [*eg.* 57, 80, 81]. In this way, the lysis they catalyze is coordinated with the biosynthesis and insertion of new PG strands (Holtje 1998). Some LTs, however, remain free within the periplasm and so another level of control is required which occurs through the O-acetylation of the PG substrate. PG O-acetylation occurs at the C-6 hydroxyl of MurNAc residues (Fig. 1), thereby precluding LT activity. Gram-negative bacteria that use this form of substrate modification to control LT activity also produce O-acetylPG esterases which remove the blocking O-acetylation as required (Weadge and Clarke 2005). Although not yet demonstrated, it has been proposed that these esterases would also associate with the PG biosynthetic complexes, presumably at their leading edges, to clear away the O-acetylation thereby permitting the localized lysis of the sacculus to provide new sites for PG insertion (Moynihan and Clarke 2011). Regardless, the level of PG O-acetylation was found to decrease from 51 % to 29 % (relative to MurNAc content) upon differentiation of *Proteus mirabilis* vegetative cells to hyper-flagellated swarmers (Strating et al. 2012). This decrease was accompanied by changes in the autolysin profile of the cells and in the muropeptide composition of the PG, changes reflective of the requirement to provide the pores for flagella insertion. However, not all bacteria O-acetylate their PG, notably including *E. coli* and *P. aeruginosa* (Scheurwater et al. 2008; Moynihan and Clarke 2011). Those that do not instead produce specific proteinaceous inhibitors

that are specifically localized to the periplasm and block LT activity through protein complexation (Clarke et al. 2010; Pfeffer et al. 2012).

4.2 Control and Modulation of Flagella-Specific Autolytic Activity

Very recently, it has been shown that specific variations of the themes described above for autolytic control have been adopted for flagellum assembly. In β- and γ-proteobacteria, the autolytic β-*N*-acetylglucosaminidase activity of FlgJ associated with its C-terminal domain is understood to be controlled by the function of its N-terminal domain (Herlihey et al. 2014). This N-terminal domain of FlgJ acts as a scaffolding rod-capping protein and thus it associates with the developing rod upon its production and thereby securing its localization at the site of flagella formation (Hirano et al. 2001; Nambu et al. 1999). After the hole has been bored in the PG sacculus by its C-terminal domain, FlgJ remains associated with the growing rod as a structural component through its N-terminal domain. Consequently, the β-*N*-acetylglucosaminidase domain of FlgJ becomes localized extracellularly (Cohen and Hughes 2014) thereby preventing any deleterious autolytic activity. In contrast, the rod-capping FlgJ of the α-proteobacteria lacks the β-*N*-acetylglucosaminidase domain and to compensate, the developing rod recruits the exogenous SltF for the necessary autolysis (de la Mora et al. 2012). In so doing, SltF is brought into proximity of two other rod proteins, FlgB and FlgF. Using *R. sphaeroides* as the model system, these two proteins were shown to bind to, and modulate, the activity of SltF (Herlihey et al. 2016). FlgB both stabilizes and enhances the activity of SltF whereas FlgF inactivates the enzyme. The production of FlgB is an early event in the assembly of the flagellum and so its association with SltF would serve to promote the localized autolysis through the PG sacculus to accommodate the growing rod. With the space created, the later produced FlgF then binds the complex to

completely block further lysis (Herlihey et al. 2016). These interactions thus ensure SltF's activity is not left uncontrolled within the periplasm as it would remain fixed to the developing rod of the flagellum.

Currently, it is not known if homologs of FlgB and FlgF in β- and γ-proteobacteria are likewise involved in controlling the autolytic activity of FlgJ. Additionally, the regulation of the LTs involved in flagellum motility is unknown. Nonetheless, the association of specialized autolytic enzymes with macromolecular complexes appears to be a common theme. Thus, in addition to the complexes that appear to form between the PBPs and LTs for PG biosynthesis (Yunck et al. 2016; Romeis and Höltje 1994; Legaree and Clarke 2008), interactions of specialized LTs with components of secretion apparti have been described for, eg., EtgA from the injectisome type III secretion system (Burkinshaw et al. 2015) and VirB from the type IV secretion system (Höpper et al. 2005). These associations, coupled with the localization of many of the autolysins to one of the two membrane surfaces facing the periplasm serves to provide stringent control to an otherwise catastrophic activity.

5 Concluding Remarks

Significant progress has been made over the past 10 years regarding the assembly of the flagellum through the bacterial cell envelope, and it is now understood that the modulation of localized areas of PG includes providing both sites for their insertion and subsequent anchoring for proper motor function. However several questions remain regarding the lytic events, such as what directs the recognition of a specific area for insertion, and does the scaffolding of the flagellum through the envelope involve the covalent attachment of one or more of its rod proteins to PG? Insight into these and other issues pertaining to the modulation of the PG sacculus for flagellum assembly may provide new opportunities to further exploit PG metabolism for the development

of new antibacterials, an issue of increasing importance is this era of multidrug resistance.

Conflict of Interest The authors declare no conflict of interest.

References

Aika T, Yoshimura H, Namba K (1990) Monolayer crystallization of flagellar L-P rings by sequential addition and depletion of lipid. Science 252:1544–1546

Artola-Recolons C, Carrasco-López C, Llarrull LI, Kumarasiri M, Lastochkin E, Martínez de Ilarduya I, Meindl K, Usón I, Mobashery S, Hermoso JA (2011) High-resolution crystal structure of an outer membrane-anchored endolytic peptidoglycan lytic transglycosylases (MltE) from *Escherichia coli*. Biochemistry 50(13):2384–2386

Artola-Recolons C, Lee M, Bernardo-García N, Blázquez B, Hesek D, Bartual SG, Mahasenan KV, Lastochkin E, Pi H, Boggess B, Meindl K, Usón I, Fischer JF, Mobashery S, Hermoso JA (2014) Structure and cell wall cleavage by modular lytic transglycosylases MltC of *Escherichia coli*. ACS Chem Biol 9:2058–2066

Asai Y, Kojima S, Kato H, Nishioka N, Kawagishi I, Homma M (1997) Putative channel components for the fast-rotating sodium-driven flagellar motor of a marine bacterium. J Bacteriol 176(16):5104–5110

Bai XH, Chen HJ, Jiang YL, Wen Z, Huang Y, Cheng W, Li Q, Qi L, Zhang JR, Chen Y, Zhou CZ (2014) Structure of pneumococcal peptidoglycan hydrolase LytB reveals insights into the bacterial cell wall remodeling and pathogenesis. J Biol Chem 289:23403–23416

Blackburn NT, Clarke AJ (2001) Identification of four families of peptidoglycan lytic transglycosylases. J Mol Evol 52:78–84

Bublitz M, Polle L, Holland C, Heinz DW, Nimtz M, Schubert W (2009) Structural basis for autoinhibition and activation of Auto, a virulence- associated peptidoglycan hydrolase of Listeria monocytogenes. Mol Microbiol 71:1509–1522

Burkinshaw BJ, Deng W, Lameignère E, Wasney GA, Zhu H, Worrall LJ, Finlay BB, Strynadka NC (2015) Structural analysis of a specialized type III secretion system peptidoglycan-cleaving enzyme. J Biol Chem 290:10406–10417

Chen R, Guttenplan SB, Blair KM, Kearns DB (2009) Role of the σ^D-dependent autolysins in *Bacillus*

subtilis population heterogeneity. J Bacteriol 191 (18):5775–5784

Clarke CA, Scheurwater EM, Clarke AJ (2010) The vertebrate lysozyme inhibitor Ivy functions to inhibit the activity of lytic transglycosylases. J Biol Chem 285:14843–14847

Cohen EJ, Hughes KT (2014) Rod-to-hook transition for extracellular flagellum assembly is catalyzed by the L-ring dependent rod scaffold removal. J Bacteriol 196:2387–2395

Davies GJ, Wilson KS, Henrissat B (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. Biochem J 321:557–559

de la Mora J, Ballado T, González-Pedrajo B, Camerena L, Dreyfus G (2007) The flagellar muramidase from the photosynthetic bacterium *Rhodobacter sphaeroides*. J Bacteriol 189:7998–8004

de la Mora J, Osorio-Valeriano M, González-Pedrajo B, Ballado T, Camerena L, Dreyfus G (2012) The C terminus of the flagellar muramidase SltF modulates the interaction with FlgJ in *Rhodobacter sphaeroides*. J Bacteriol 194:4513–4520

De Mot R, Vanderleyden J (1994) The C-terminal sequence conservation between OmpA-related outer membrane proteins and MotB suggests a common function in both Gram-positive and Gram-negative bacteria, possibly in the interaction of these domains with peptidoglycan. Mol Microbiol 12(2):333–334

Demchick P, Koch AL (1996) The permeability of the wall fabric of *Escherichia coli* and *Bacillus substilis*. J Bacteriol 178(3):768–773

Drouillard S, Armand S, Davies GJ, Vorgias CE, Henrissat B (1997) *Serratia marcescens* chitobiase is a retaining glycosidase utilizing substrate acetamido group participation. Biochem J 328:945–949

González-Pedrajo B, de la Mora J, Ballado T, Camerena L, Dreyfus G (2002) Characterization of the *flgG* operon of *Rhodobacter sphaeroides* WS8 and its role in flagellum biosynthesis. Biochim Biophys Acta 1579:55–63

Gumbart JC, Beeby M, Jensen GJ, Roux B (2014) *Escherichia coli* peptidoglycan structure and mechanics as predicted by atomic-scale stimulations. PLoS Comput Biol 10(2):e1003475. doi:10.1371/journal. pcbi.1003475

Hashimoto W, Ochiai A, Momma K, Itoh T, Mikami B, Maruyama Y, Murata K (2009) Crystal structure of the glycosidase family 73 peptidoglycan hydrolase FlgJ. Biochem Biophys Res Commun 381(1):16–21

Herlihey FA, Moynihan PJ, Clarke AJ (2014) The essential protein for bacterial flagella formation FlgJ functions as a β-*N*-acetylglucosaminidase. J Biol Chem 289(45):31029–31042

Herlihey FA, Osorio-Valeriano M, Dreyfus G, Clarke AJ (2016) Modulation of the lytic activity of the dedicated autolysin for flagella formation SltF by flagella rod proteins FlgB and FlgF. J Bacteriol. doi:10.1128/ JB.00203-16

Hirano T, Minamino T, Macnab RM (2001) The role in flagellar rod assembly of the N-terminal domain of *Salmonella* FlgJ, a flagellum-specific muramidase. J Mol Biol 312:359–369

Hizukuri Y, Morton JF, Yakushi T, Kojima S, Homma M (2009) The peptidoglycan-binding (PGB) domain of the *Escherichia coli* Pal protein can also function as the PGB domain in *E. coli* flagellar motor protein MotB. J Biochem 146:219–229

Holtje J-V (1998) Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*. Microbiol Mol Biol Rev 62:181–203

Höltje JV, Mirelman D, Sharon N, Schwarz U (1975) Novel type of murein transglycosylases in *Escherichia coli*. J Bacteriol 124:1067–1076

Homma M, Komeda Y, Iino T, Macnab RM (1987) The *flaFIX* gene product of *Salmonella typhimurium* is a flagellar basal body component with a signal peptide for export. J Bacteriol 169:1493–1498

Höpper C, Carle A, Sivanesan D, Hoeppner S, Baron C (2005) The putative lytic transglycosylases VirB from *Brucella suis* interacts with the type IV secretion system core components VirB8, VirB9 and VirB11. Microbiology 151:3469–3482

Hoskin ER, Vogt C, Bakker EP, Manson MD (2006) The *Escherichia coli* MotAB proton channel unplugged. J Mol Biol 364:921–937

Inagaki N, Iguchi A, Yokoyama T, Yokoi KJ, Ono Y, Yamakawa A, Taketo A, Kodaira K (2009) Molecular properties of the glucosaminidase AcmA from Lactococcus lactis MG1363: mutational and biochemical analyses. Gene 447(2):61–71

Jorgenson MA, Chen Y, Yahashiri A, Popham DL, Weiss DS (2014) The bacterial septal ring protein RlpA is a lytic transglycosylases that contributes to rod shape and daughter cell separation in *Pseudomonas aeruginosa*. Mol Microbiol 93(1):113–128

Knapp S, Vocadlo DJ, Gao Z, Kirk B, Lou J, Withers SG (1996) NAG-thiazoline, An N-Acetyl-β-hexosaminidase inhibitor that implicates acetamido participation. J Am Chem Soc 118:6804–6805

Kojima S, Blair DF (2001) Conformational change in the stator of the bacterial flagellar motor. Biochemistry 40:13041–13050

Kojima S, Imada K, Sakuma M, Sudo Y, Kojima C, Minamino T, Homma M, Namba K (2008a) Stator assembly and activation mechanism of the flagellar motor by the periplasmic region of MotB. Mol Microbiol 73:710–718

Kojima S, Shinohara A, Terashima H, Yakushi T, Sakuma M, Homma M, Namba K, Imada K (2008b) Insights into the stator assembly of the *Vibrio* flagellar motor from the crystal structure of MotY. Proc Natl Acad Sci U S A 105(22):7696–7701

Koushik P, Erhardt M, Hirano T, Blair DF, Hughes KT (2008) Energy source of flagellar type III secretion. Nature 451:489–492

Kubori T, Shimamoto N, Shigeru Y, Namba K, Aizawa S-I (1992) Morphological pathway of flagellar assembly in *Salmonella typhimurium*. J Mol Biol 226:433–446

Lee M, Hesek D, Llarrull LI, Lastochkin E, Pi H, Boggess B, Mobashery S (2013) Reactions of all *E. coli* lytic transglycosylases with bacterial cell wall. J Am Chem Soc 135(9):3311–3314

Lee M, Domínguez-Gil T, Hesek D, Mahasenan KV, Lastochkin E, Hermoso JA, Mobashery S (2016) Turnover of bacterial cell wall by SltB3, a multidomain lytic transglycosylases of *Pseudomonas aeruginosa*. ACS Chem Biol. doi:10.1021/acschembio.6b00194

Legaree BA, Clarke AJ (2008) Interaction of penicillin-binding protein 2 with soluble lytic transglycosylases B1 in *Pseudomonas aeruginosa*. J Bacteriol 190(20):6922–6926

Leung AKW, Duewel HS, Honek JF, Berghuis AM (2001) Crystal structure of the lytic transglycosylases from bacteriophage lambda in complex with hexa-*N*-acetylchitohexaose. Biochemistry 40:5665–5673

Lipski A, Hervé M, Lombard V, Nurizzo D, Mengin-Lecreulx D, Bourne Y, Vincent F (2015) Structural and biochemical characterization of the β-*N*-acetylglucosaminidase from *Thermotoga maritima*: toward rationalization of mechanistic knowledge in the GH73 family. Glycobiology 25(3):319–330

Lombard V, Golaconda Ramulu H, Drula E, Countinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490–D495

Macdonald JM, Tarling CA, Taylor EJ, Dennis RJ, Myers DS, Knapp S, Davies GJ, Withers SG (2010) Chitinase inhibition by chitobiose and chitotriose thiazolines. Angew Chem Int Ed Engl 49(14):2599–2602

Macnab RM (2003) How bacteria assemble flagella. Annu Rev Microbiol 577:77–100

Maruyama Y, Ochiai A, Itoh T, Mikami B, Hashimoto W, Murata K (2010) Mutational studies of the peptidoglycan hydrolase FlgJ of Sphingomonas sp. strain A1. J Basic Microbiol 50(4):311–317

McCarter JD, Withers SG (1994) Mechanisms of enzymatic glycoside hydrolysis. Curr Opin Struct Biol 4:885–892

Minamino T, Imada K (2015) The bacterial flagellar motor and its structural diversity. Trends Microbiol 23(5):267–274

Minamino T, Macnab RM (1999) Components of the *Salmonella* flagellar export apparatus and classification of export substrates. J Bacteriol 181:1388–1394

Minaminoa T, Namba K (2008) Distinct roles of the FliI ATPase and proton motive force in bacterial flagellar protein export. Nature 451:485–488

Moynihan PJ, Clarke AJ (2011) O-Acetylated peptidoglycan: controlling the activity of bacterial autolysins and lytic enzymes of innate immune systems. Int J Biochem Cell Biol 43:1655–1659

Nambu T, Minamino T, Macnab RM, Kutsukake K (1999) Peptidoglycan-hydrolyzing activity of the FlgJ protein, essential for flagellar rod formation in *Salmonella typhimurium*. J Bacteriol 181:1555–1561

Nambu T, Inagaki Y, Kutsukake K (2006) Plasticity of the domain structure in FlgJ, a bacterial protein involved in flagellar rod formation. Genes Genet Syst 81:381–389

Nikolaidis I, Izore T, Job V, Thielens N, Breukink E, Dessen A (2012) Calcium-dependent complex formation between PBP2 and lytic transglycosylases SltB1. Microb Drug Resist 18:298–305

Osorio-Valeriano M, de la Mora J, Camarena L, Dreyfus G (2016) Biochemical characterization of the flagellar rod components of *Rhodobacter sphaeroides*: properties and interactions. J Bacteriol 198 (3):544–552

Park JS, Lee WC, Yeo KJ, Ryu K-S, Kumarasiri M, Hesek D, Lee M, Mobashery S, Song JH, Kim SI, Lee JC, Cheong C, Jeon YH, Kim H-Y (2012) Mechanism of anchoring of OmpA protein to the cell wall peptidoglycan of the gram-negative bacterial outer membrane. FASEB J 26(1):219–228

Parsons LM, Lin F, Orban J (2006) Peptidoglycan recognition by Pal, an outer membrane lipoprotein. Biochemistry 45:2122–2128

Pfeffer JM, Moynihan PJ, Clarke CA, Clarke AJ (2012) Control of lytic transglycosylase activity within bacterial cell walls. In: Reid, Twine, Read (eds) Bacterial glycomics. Caister Academic Press, Norfolk, pp 55–68

Powell AJ, Liu ZJ, Nicholas RA, Davies C (2006) Crystal structure of the lytic transglycosylases MltA from *N. gonorrhoeae* and *E. coli*: insights into interdomain movements and substrate binding. J Mol Biol 359:122–136

Reboul CF, Andrews DA, Nahar MF, Buckle AM, Roujeinikova A (2011) Crystallographic and molecular dynamics analysis of loop motions unmasking the peptidoglycan-binding site in stator protein MotB of flagellar motor. PLoS One 6(4):e18981. doi:10.1371/journal.pone.0018981

Reid CW, Blackburn NT, Legaree BA, Auzanneau F-I, Clarke AJ (2004) Inhibition of membrane-bound lytic transglycosylase B by NAG-thiazoline. FEBS Lett 574:73–79

Reid CW, Legaree BA, Clarke AJ (2007) Role of Ser216 in the mechanism of action of membrane-bound lytic transglycosylases B: further evidence for substrate-assisted catalysis. FEBS Lett 581:4988–4992

Rico-Lastres P, Díez-Martínez R, Iglesias-Bexiga M, Bustamante N, Aldridge C, Hesek D, Lee M, Mobashery S, Gray J, Vollmer W, Garcia P, Menéndez M (2015) Substrate recognition and catalysis by LytB, a pneumococcal peptidoglycan hydrolase involved in virulence. Sci Rep 5:16198

Romeis T, Höltje JV (1994) Specific interaction of penicillin-binding proteins 3 and 7/8 with soluble lytic transglycosylases in Escherichia coli. J Biol Chem 269(34):21603–21607

Roujeinikova A (2008) Crystal structure of the cell wall anchor domain of MotB, a stator component of the bacterial flagellar motor: implications for peptidoglycan recognition. Proc Natl Acad Sci U S A 105:10348–10353

Roure S, Bonis M, Chaput C, Ecobichon C, Mattox A, Barrière C, Geldmacher N, Guadagnini S, Schmitt C, Prèvost MC, Labigne A, Backert S, Ferrero RL, Boneca IG (2012) Peptidoglycan maturation enzymes affect flagellar functionality in bacteria. Mol Micrbiol 86:845–856

Scheurwater EM, Clarke AJ (2008) The C-terminal domain of *Escherichia coli* YhfD functions as a lytic transglycosylases. J Biol Chem 283(13):8363–8373

Scheurwater E, Reid CW, Clarke AJ (2008) Lytic transglycosylases: bacterial space-making autolyins. Int J Biochem Cell Biol 40:586–591

Strating H, Vandenende C, Clarke AJ (2012) Changes in peptidoglycan structure and metabolism during differentiation of *Proteus mirabilis* into swarmer cells. Can J Microbiol 58:1183–1194

Sudiarta IP, Fukushima T, Sekiguchi J (2010) Bacillus subtilis CwlQ (previously YjbJ) is a bifunctional enzyme exhibiting muramidase and soluble-lytic transglycosylases activities. Biochem Biophys Rec Commun 398:606–612

Suzuki H, Yonekura K, Murata K, Hirai T, Oosawa K, Namba K (1998) A structural feature in the central channel of the bacterial flagellar FliE ring complex is implicated in type III protein export. J Struct Biol 124:104–114

Terashima H, Fukuoka H, Yakushi T, Kojima S, Homma M (2006) The *Vibrio* motor proteins, MotX and MotY, are associated with the basal body of Na$^+$-driven flagella and required for stator formation. Mol Microbiol 62(4):1170–1180

Terashima H, Koike M, Kojima S, Homma M (2010) The flagellar basal body-associated protein FlgT is essential for a novel ring structure in the sodium-driven *Vibrio* motor. J Bacteriol 192(21):5609–5615

Terwissscha van Scheltinga AC, Armand S, Kalk KH, Isogai A, Henrissat B, Dijkstra BW (1995) Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and x-ray structure of a complex with allosamidin: evidence for substrate assisted catalysis. Biochemistry 34:15619–15623

Tews I, Perrakis A, Oppenheim A, Dauter Z, Wilson KS, Vorgias CE (1996) Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. Nat Struct Biol 3:638–648

Thunnissen A-MWH, Isaacs NW, Dijkstra BW (1995) The catalytic domain of a bacterial lytic transglycosylases defines a novel class of lysozymes. Proteins 22:245–258

van Asselt EJ, Dijkstra AJ, Kalk KH, Takacs B, Keck W, Dijkstra BW (1999a) Crystal structure of *Escherichia*

coli lytic transglycosylases Slt35 reveals a lysozyme-like catalytic domain with an EF-hand. Structure 7:1167–1180

van Asselt EJ, Thunnissen A-MWH, Dijkstra BW (1999b) High resolution crystal structures of the *Escherichia coli* lytic transglycosylases Slt70 and its complex with a peptidoglycan fragment. J Mol Biol 291:877–898

van Asselt EJ, Kalk KH, Dijkstra BW (2000) Crystallographic studies of the interactions of *Escherichia coli* lytic transglycosylases Slt35 with peptidoglycan. Biochemistry 39:1924–1934

van Straaten KE, Dijkstra BW, Vollmer W, Thunnissen A-MWH (2005) Crystal structure of MltA from *Escherichia coli* reveals a unique lytic transglycosylases fold. J Mol Biol 353:1068–1082

Vázquez-Laslop N, Hyunwoo L, Hu R, Neyfakh AA (2001) Molecular sieve mechanism of selective release of cytoplasmic proteins by osmotically shocked *Escherichia coli*. J Bacteriol 183 (8):2399–2404

Viollier PH, Shapiro L (2003) A lytic transglycosylases homologue, PleA, is required for the assembly of pili and the flagellum at the *Caulobacter crescentus* cell pole. Mol Microbiol 49:331–345

Vollmer W, Bertsche U (2008) Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. Biochim Biophys Acta 1178:1714–1734

Vollmer W, Joris B, Charlier P, Foster S (2008) Bacterial peptidoglycan (murein) hydrolases. FEMS Microbiol Rev 32:259–286

Weadge JT, Clarke AJ (2005) Identification and characterization of *O*-acetylpeptidoglycan esterase: a novel enzyme discovered in *Neisseria gonorrhoeae*. Biochemistry 45:839–851

Yokoi KJ, Sugahara K, Iguchi A, Nishitani G, Ikeda M, Shimada T, Inagaki N, Yamakawa A, Taketo A, Kodaira K (2008) Molecular properties of the putative autolysin Atl(WM) encoded by Staphylococcus warneri M: mutational and biochemical analyses of the amidase and glucosaminidase domains. Gene 416 (1–2):66–76

Yunck R, Hongbaek C, Bernhardt TG (2016) Identification of MltG as a potential terminase for peptidoglycan polyermization in bacteria. Mol Microbiol 99 (4):700–718

Zabola P, Bailey-Elkin BA, Derksen M, Mark BL (2016) Structural and biochemical insights into the peptidoglycan hydrolase domain of FlgJ from *Salmonell typhimurium*. PLoS One 11(2):e0149204

Zhao X, Zhang K, Boquoi T, Hu B, Motaleb MA, Miller KA, James ME, Charon NW, Manson MD, Norris SJ, Li C, Liu J (2013) Cryoelectron tomography reveals the sequential assembly of bacterial flagellum in *Borrelia burgdorferi*. Proc Natl Acad Sci U S A 110 (35):14390–14395

Zhu S, Takao M, Li N, Sakuma M, Nishino Y, Homma M, Kojima S, Imada K (2014) Conformational change in the periplasmic region of the flagellar stator coupled with the assembly around the rotor. Proc Natl Acad Sci U S A 111(37):13523–13528

Regulation of Skeletal Muscle Myoblast Differentiation and Proliferation by Pannexins

Stéphanie Langlois and Kyle N. Cowan

Abstract

Pannexins are newly discovered channels that are now recognized as mediators of adenosine triphosphate release from several cell types allowing communication with the extracellular environment. Pannexins have been associated with various physiological and pathological processes including apoptosis, inflammation, and cancer. However, it is only recently that our work has unveiled a role for Pannexin 1 and Pannexin 3 as novel regulators of skeletal muscle myoblast proliferation and differentiation. Myoblast differentiation is an ordered multistep process that includes withdrawal from the cell cycle and the expression of key myogenic factors leading to myoblast differentiation and fusion into multinucleated myotubes. Eventually, myotubes will give rise to the diverse muscle fiber types that build the complex skeletal muscle architecture essential for body movement, postural behavior, and breathing. Skeletal muscle cell proliferation and differentiation are crucial processes required for proper skeletal muscle development during embryogenesis, as well as for the postnatal skeletal muscle regeneration that is necessary for muscle repair after injury or exercise. However, defects in skeletal muscle cell differentiation and/or deregulation of cell proliferation are involved in various skeletal muscle pathologies. In this review, we will discuss the expression of pannexins and their post-translational modifications in skeletal muscle, their known functions in various steps of myogenesis,

K.N. Cowan (✉)
Department of Surgery, Division of Surgery, University of Ottawa, Children's Hospital of Eastern Ontario, 401 Smyth Road, Room 3370, Ottawa, ON K1H 8 L1, Canada

Molecular Biomedicine Program, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada

Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada
e-mail: kcowan@cheo.on.ca

S. Langlois
Department of Surgery, Division of Surgery, University of Ottawa, Children's Hospital of Eastern Ontario, 401 Smyth Road, Ottawa, ON K1H 8 L1, Canada

Molecular Biomedicine Program, Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada

including myoblast proliferation and differentiation, as well as their possible roles in skeletal muscle development, regeneration, and diseases such as Duchenne muscular dystrophy.

Keywords

Pannexin • Skeletal muscle • Myoblast • Proliferation • Differentiation

Abbreviations

ATP	Adenosine triphosphate
Ca^{2+}	Calcium
CBX	Carbenoxolone
DMD	Duchenne muscular dystrophy
HSMM	Human primary skeletal muscle myoblast
kDa	Kilodalton
P2R	P2 receptor
Panx	Pannexin
Panx1	Pannexin 1
Panx2	Pannexin 2
Panx3	Pannexin 3
RC	Reserve cell
RMS	Rhabdomyosarcoma
SC	Satellite cell

1   Pannexins

Pannexins (Panxs) are mammalian orthologs of the invertebrate gap junction proteins, innexins, that have been discovered about 15 years ago (Panchin et al. 2000; Baranova et al. 2004). The Panx family is composed of three members in the mammalian genome: Panx1, Panx2, and Panx3 (Panchin et al. 2000; Baranova et al. 2004). Panxs have a membrane topology consisting of four membrane-spanning domains, two extracellular loops, a cytoplasmic loop, and intracellular N- and C-termini (Panchin 2005) and function as single membrane channels (Sosinsky et al. 2011; Beckmann et al. 2016). Panx channels have been mainly involved in controlling the exchange of small molecules and ions between the cytosol and extracellular space such as adenosine triphosphate (ATP) release into the extracellular milieu (Bao et al. 2004; Huang et al. 2007;

Chekeni et al. 2010; Locovei et al. 2006a; Iwamoto et al. 2010; Pelegrin and Surprenant 2006; Qu et al. 2011; Suadicani et al. 2012; Timoteo et al. 2014; Dolmatova et al. 2012; Riquelme et al. 2013; Pinheiro et al. 2013). Some studies have also suggested that Panxs can function as calcium (Ca^{2+}) channels in the endoplasmic reticulum (D'Hondt et al. 2011; Vanden Abeele et al. 2006; Ishikawa et al. 2011). In addition, Panx1 has been suggested to be involved in glucose uptake in electrically stimulated skeletal muscle fibers (Riquelme et al. 2013).

Panx1 is ubiquitously expressed in many organs and tissues. Several studies have reported that human *PANX2* and murine *Panx2* transcripts are mainly restricted to the central nervous system (Baranova et al. 2004; Bruzzone et al. 2003; Zoidl et al. 2008; Bond et al. 2012). However, it was recently demonstrated that *Panx2* transcriptional activity does not correlate well with Panx2 protein levels (Le Vasseur et al. 2014). As opposed to that initially predicted based on transcript levels, Panx2 seems to have a ubiquitous protein expression profile (Le Vasseur et al. 2014). On the other hand, Panx3 is most abundant in skin, bone, and cartilage (Shestopalov and Panchin 2008). Panx3, with an expected molecular weight of ~43–44 kilodaltons (kDa), exhibits an immunoreactive species at ~70 kDa in several human and rodent tissues and cell types (Penuela et al. 2007, 2014a; Celetti et al. 2010; Cowan et al. 2012; Turmel et al. 2011; Langlois et al. 2014). While the exact identity of the ~70 kDa Panx3 immunoreactive species remains unknown, it likely does not correspond to a Panx3 dimer (Langlois et al. 2014) but to a glycoprotein that is recognized by three different antibodies directed

against Panx3 (Penuela et al. 2007; Cowan et al. 2012). Our recent data demonstrated that the levels of the ~70 kDa immunoreactive species expressed in human primary skeletal muscle myoblasts (HSMM) can be reduced using short hairpin RNAs (shRNAs) against *PANX3* (Langlois et al. 2014). While these results collectively suggest that this ~70 kDa immunoreactive band corresponds to a Panx3 species, precaution should be taken when referring to it until such time as its sequence and its exact identity elucidated.

Amongst Panx members, Panx1 has been the most studied thus far. Panx1 channels have been implicated in many cellular and physiological functions such as long range Ca^{2+} wave propagation (Locovei et al. 2006b), vasodilatation (Locovei et al. 2006a), inflammatory responses (Pelegrin and Surprenant 2006; Silverman et al. 2009), leukocyte emigration through the venous endothelium during acute inflammation (Lohman et al. 2015), human immunodeficiency virus infection (Orellana et al. 2013; Seror et al. 2011; Paoletti et al. 2013), neuronal cell death (Thompson et al. 2006; Weilinger et al. 2016; Gulbransen and Sharkey 2012), epilepsy (Thompson et al. 2008), apoptosis (Chekeni et al. 2010; Qu et al. 2011), stabilization of synaptic plasticity and learning (Prochnow et al. 2012), keratinocyte differentiation (Celetti et al. 2010), and carcinogenesis (Cowan et al. 2012; Lai et al. 2007; Lai et al. 2009; Penuela et al. 2013). By contrast to Panx1, much less is known about the cellular and physiological functions of Panx2 and Panx3. Panx2 has been implicated in neuronal commitment (Swayne et al. 2010) and in the regulation of C6 glioma cell growth (Lai et al. 2009). Together with Panx1, Panx2 channel function was shown to contribute to ischemic brain damage (Bargiotas et al. 2011). As for Panx3, we have recently reported that its levels are altered in human keratinocyte tumors suggesting a role in carcinogenesis (Cowan et al. 2012). Studies have also reported a role for Panx3 in regulating chondrocytes (Iwamoto et al. 2010) and osteoprogenitor cell proliferation (Ishikawa et al. 2014), as well as in promoting osteoblast

(Ishikawa et al. 2011) and chondrocyte (Iwamoto et al. 2010) differentiation. Using *Panx3* knock out mice, it has been recently demonstrated that Panx3 is required for normal progression of skeletal development (Oh et al. 2015; Caskenette et al. 2016; Ishikawa et al. 2016) and that mice lacking *Panx3* are resistant to the development of osteoarthritis (Moon et al. 2015).

2 Pannexin Expression in Skeletal Muscle

Only recently have pannexins been studied in the skeletal muscle. *PANX1* transcripts were found to be highly expressed in human skeletal muscle (Baranova et al. 2004) and Panx1 protein has been detected by western blotting in mouse, rat, and human skeletal muscles (Riquelme et al. 2013; Langlois et al. 2014). In addition, differentiated myoblasts (Langlois et al. 2014), myotubes (Buvinic et al. 2009), and myofibers all express Panx1 (Riquelme et al. 2013; Cea et al. 2013). Panx1 antibodies recognize several species (mainly ~38–50 kDa) in human, mouse, and rat skeletal muscle tissue, most likely reflecting various degrees of glycosylation (Langlois et al. 2014). Depending of the tissue examined for endogenous Panx expression, immunolabeling of Panx1 has revealed a punctate pattern, such as in human epidermis (Cowan et al. 2012) and mouse mammary gland (Stewart et al. 2016), as well as diffuse staining as seen, for example, in human hair follicle (Cowan et al. 2012) and in several cell types within the human colon (Diezmos et al. 2013). In longitudinal sections of human and rat skeletal muscles, Panx1 was detected as punctate staining (Fig. 1) (Riquelme et al. 2013; Langlois et al. 2014). However, in transverse sections of rat gastrocnemius, the Panx1 labeling, which was present in the interior of the fibers, was more diffuse (Riquelme et al. 2013). More specifically, Panx1 protein has been localized at the sarcolemma (Jorquera et al. 2012) and T-tubules of adult mice and rat skeletal muscle fibers (Riquelme et al. 2013; Jorquera et al. 2012).
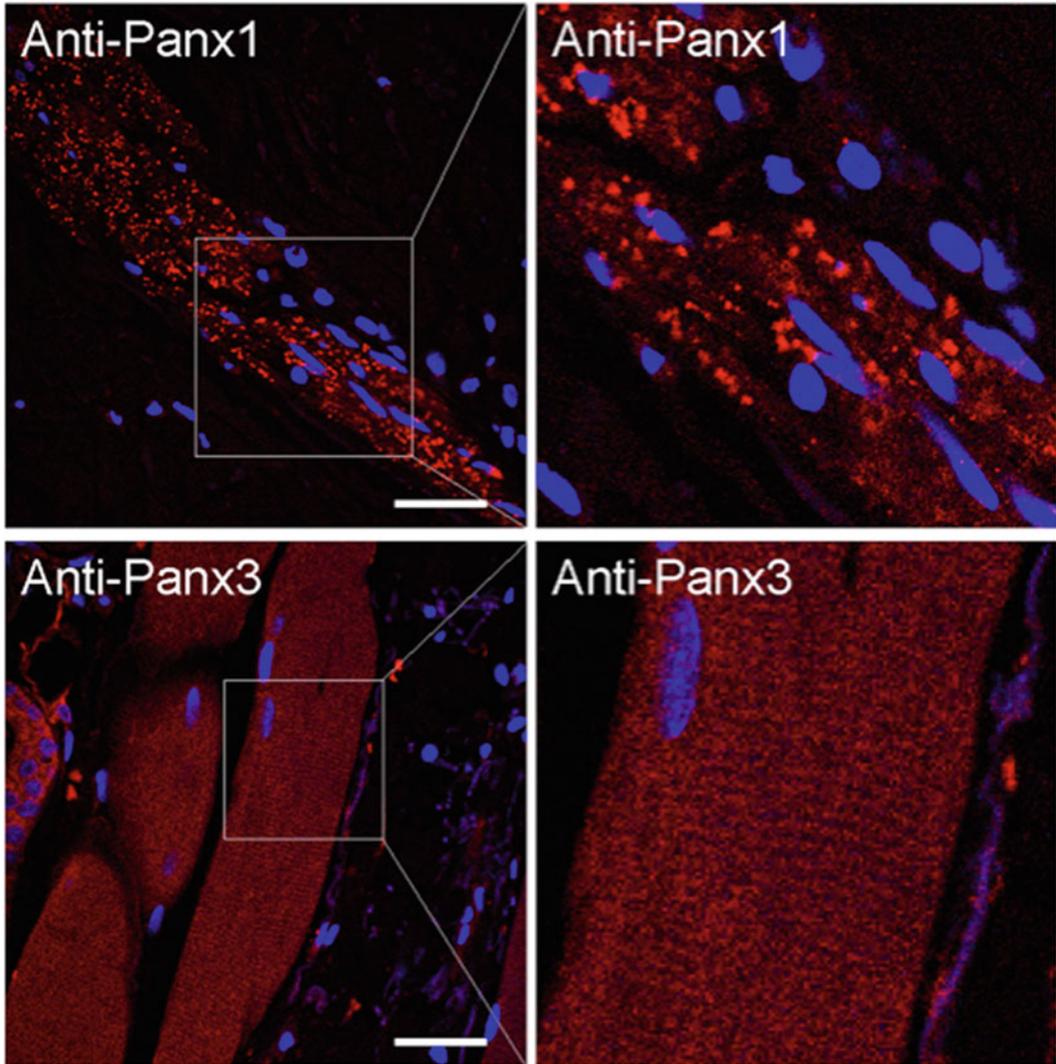
Fig. 1 PANX1 and PANX3 show a different labeling pattern in human skeletal muscle tissue.
Representative images of human skeletal muscle tissue in skin samples labeled for PANX1 (labeled in *red*) and PANX3 (labeled in *red*). PANX1 was detected as a punctate stain, whereas PANX3 was observed as diffuse labeling. Higher magnification micrographs of PANX3 labeling show a striated pattern. *Blue* = nuclei; *bars* = 50 μm (This research was originally published in Ref. (Langlois et al. 2014). © the American Society for Biochemistry and Molecular Biology)

Riquelme et al. could not detect Panx2 in several rat skeletal muscle types (extensor digitorum longus, soleus, and flexor digotorum brevis) by western blotting using commercially available Panx2 antibodies developed in rabbits, while it was present as expected in the brain (Riquelme et al. 2013). In our hands, Panx2 was also absent or below detectable levels in human, rat, and mice skeletal muscle homogenates when assessed by western blotting using Panx2 antibodies developed by Penuela et al. which specifically recognize mouse Panx2 over-expressed in HEK293T cells (Langlois et al. 2014; Penuela et al. 2014b). However, a study published afterwards using a novel commercial monoclonal antibody (clone N121A/1) that detects mouse, rat, and human Panx2 revealed that Panx2 protein is expressed in

murine skeletal muscle tissue (Le Vasseur et al. 2014). This group further detected *Panx2* transcripts in mouse skeletal muscle by real-time qPCR (Le Vasseur et al. 2014). Pillon et al. also reported that *Panx2* is expressed in mouse quadriceps muscle, as well as *PANX2* in primary human myotubes (Pillon et al. 2014). In the light of these studies, it would be interesting and relevant to revisit the current work on pannexin function in skeletal muscle now taking into account the presence of Panx2 and its potential impact.

In addition to Panx1, we have found that Panx3 protein is also expressed in mouse, rat, and human skeletal muscle (Langlois et al. 2014). Bands at the expected molecular weight of Panx3 at about 43 kDa were detected, as well as its ~70 kDa immunoreactive species (Langlois et al. 2014). As opposed to the punctate PANX1 staining, PANX3 was detected as a diffuse pattern of labeling in human skeletal muscle tissue (Fig. 1) (Langlois et al. 2014). This staining likely corresponds to that of the ~43 kDa species as the antibodies used previously did not recognize the ~70 kDa immunoreactive species expressed in rat epidermal keratinocytes by immunofluorescence microscopy (Celetti et al. 2010).

Given that Panx1 and Panx3 and their roles in skeletal muscle have been the only ones investigated so far, the reminder of this review will focus primarily on Panx1 and Panx3 channels.

3 Post-translational Modifications of Panx1 and Panx3

3.1 Glycosylation

Panx1 and Panx3 are known to be heavily *N*-glycosylated, generating distinct species in rat, mice, and human cells and tissues including skeletal muscle (Riquelme et al. 2013; Penuela et al. 2007; Langlois et al. 2014; Penuela et al. 2009; Boassa et al. 2007). Using site-directed mutagenesis to inhibit *N*-glycosylation,

it has been clearly demonstrated that Panx1 is glycosylated on the second extracellular loop at position Asn-254 (N254) and that this post-translational modification is important for its plasma membrane targeting (Penuela et al. 2007; Boassa et al. 2007). As a result of their different degrees of glycosylation, various molecular weight forms of Panx1 and Panx3 are detected by western blotting. Gly0 corresponds to the unglycosylated Panx core, while Gly1 represents the high mannose species that is mainly found in the endoplasmic reticulum. Gly2 correlates to a complex glycosylated form and is the main species expressed at the cell surface (Penuela et al. 2007; 2009; Boassa et al. 2007).

Mouse, rat, and human skeletal muscle express the Gly0, Gly1, and Gly2 forms of Panx1, as well as higher molecular weight species of approximately 50 kDa (Langlois et al. 2014). While rodent skeletal muscle mainly express the Gly0, Gly1, and Gly2 forms of Panx1, human adult and fetal skeletal muscle tissues predominantly express a ~50 kDa PANX1 species (Langlois et al. 2014). This form is also the main PANX1 species expressed in differentiated HSMM and corresponds to a glycosylated form given that treatment with a mixture of deglycosylation enzymes (*N*-glycosidase F, *O*-glycanase, neuraminidase, ß(1,4)-galactosidase, and ß-*N*-acetylglucosaminidase) enabled its further migration to ~38 kDa on western blot (Langlois et al. 2014). As opposed to the Gly1 and Gly2 forms of Panx1, the ~50 kDa species did not completely disappear after deglycosylation treatment (Langlois et al. 2014; Penuela et al. 2009) suggesting potential additional post-translational modifications in skeletal muscle.

As for Panx3, site-directed mutagenesis studies have revealed that it is glycosylated at Asn-71 (N71) in the first extracellular loop (Penuela et al. 2007). Adult mouse, rat, and human skeletal muscle mainly express the Gly1 and Gly2 forms of Panx3 (Langlois et al. 2014). However, a ~51 kDa PANX3 species was also strongly detected in human skeletal muscle tissue, while its ~70 kDa immunoreactive species was present

at lower levels (Langlois et al. 2014). By contrast, the ~70 kDa immunoreactive species was the main form expressed in undifferentiated HSMM (Langlois et al. 2014). These ~51 and ~70 kDa also correspond to glycosylated species as deglycosylation treatment affected their gel mobility when assessed by electrophoresis (Langlois et al. 2014). While the exact identity of the ~70 kDa immunoreactive species of Panx3 remains to be elucidated, treatment of HSMM lysates with *N*-glycosidase F and sialidase indicated that it is modified by *N*-glycosylation and sialylation (Langlois et al. 2014). However, the gel mobility of the ~43 kDa species of mouse Panx3 as over-expressed in HEK293T cells did not seem to be affected by sialidase A treatment (Penuela et al. 2014b). This suggests that the ~70 kDa immunoreactive species of Panx3, and not the ~43 kDa forms, contain sialic acids as terminal entities of its glycan structure.

Although the role of sialylation is unknown in the context of Panxs, sialic acid residues have been shown to directly modulate voltage-dependent channel gating (Schwetz et al. 2011; Ednie and Bennett 2012). On the other hand, the role of glycosylation in regulating Panx function has been investigated and is reported to be involved in their proper trafficking, cellular localization, and intermixing, all of which can potentially regulate their channel function (Penuela et al. 2007; 2009; Boassa et al. 2007).

3.2   Phosphorylation

Immunoprecipitation experiments have suggested that endogenous rat Panx1 is phosphorylated on serine and threonine residues in resting extensor digitorum longus and soleus muscles, which was increased after electrical stimulation (Riquelme et al. 2013). However, Riquelme et al. could not detect phosphorylation on tyrosine residues in Panx1 immunoprecipitates (Riquelme et al. 2013). More recent studies using a murine neuroblastoma cell line (N2a) and human umbilical vein endothelial cells have indeed demonstrated Src-dependent phosphorylation of rat Panx1 on tyrosine residue

308 (C-terminus) and of human PANX1 on tyrosine residue 198 (intracellular loop), respectively (Lohman et al. 2015; Weilinger et al. 2016). As the pY198Panx1 antibody recognizes a human PANX1 species of ~55 kDa by western blotting, the band at ~54 kDa detected in a relatively small amount in human skeletal muscle may correspond to a phosphorylated species (Langlois et al. 2014). While it has not yet been directly assessed, phosphorylation of Panx1 channels has been suggested to mediate their channel opening (Riquelme et al. 2013; Lohman et al. 2015; Weilinger et al. 2016).

The gel mobility of the ~43 kDa species of Panx3 was unaffected by shrimp alkaline phosphatase or calf intestine phosphatase treatment (Penuela et al. 2007; Langlois et al. 2014). However, the ~70 kDa immunoreactive species of Panx3 endogenously expressed in HEK293T cells and HSMM migrated further after treatment with calf intestine phosphatase suggesting that this form is phosphorylated (Langlois et al. 2014). While treatment with phosphatase is a crude assay to tease out protein phosphorylation, this data suggests that the ~62 kDa species of PANX3 detected in fetal human skeletal muscle may correspond to a dephosphorylated form of the ~70 kDa immunoreactive species (Langlois et al. 2014). In this respect, we have shown that the levels of the various molecular weight species of human PANX1 and PANX3 are highly modulated between fetal and adult skeletal muscle tissue homogenates, suggesting an important regulation in their post-translational modifications, such as glycosylation and phosphorylation, during the development and maturation of skeletal muscle (Langlois et al. 2014).

3.3   Caspase Cleavage

Human PANX1 is a substrate for caspases 3 and 7, and a specific caspase-cleavage site within PANX1 was shown to be essential for its channel activation during apoptosis of immune cells (Chekeni et al. 2010). Using an *in vitro* cell-free caspase assay, Penuela et al. demonstrated that the murine isoform of Panx1, but not Panx3, can

be cleaved by caspases 3 and 7 and that this process was inhibited by the addition of a general caspase inhibitor (Z-VAD) (Penuela et al. 2014b). However, the possibility of Panx1 cleavage by caspases in normal, challenged, or diseased skeletal muscle cells has yet to be explored.

3.4 S-Nitrosylation

Panx1 can also be modified by S-nitrosylation at Cys-40 and Cys-346, which results in an inhibition of Panx1 current and ATP release (Lohman et al. 2012). S-nitrosylation of Panx3 over-expressed in HEK293T cells was also detected in response to S-nitrosoglutathione treatment (Penuela et al. 2014b), however the effect of this modification on Panx3 channel function has yet to be assessed. In addition, whether Panx1 and Panx3 are S-nitrosylated in skeletal muscle and what affect on their function this may have also remains uninvestigated.

4 Skeletal Muscle Myogenesis and Regeneration

Skeletal muscles within both the adult trunk and limbs develop from embryonic somites (Christ and Ordahl 1995; Scaal and Christ 2004). Somites are transient mesodermal units forming in a cranio-caudal succession by segmentation of the paraxial mesoderm on both sides of the neural tube (Biressi et al. 2007). Each newly formed somite rapidly differentiates into a ventral sclerotome and a dorsal dermomyotome from which myogenic precursors originate (Biressi et al. 2007). These precursor cells will eventually give rise to the majority of the skeletal muscles. Cells from the dermomyotome have high expression of the paired box transcription factors *Pax3* and *Pax7* and a low expression of the myogenic regulator *Myf5* (Jostes et al. 1990; Goulding et al. 1991; Kiefer and Hauschka 2001; Bentzinger et al. 2012). During the first stage of embryonic development (at embryonic day E8.75 in the mouse), a group of postmitotic

mononucleated myocytes migrate out from the border regions of the dermomyotome and form primitive muscles constituting the primary myotome beneath the dermomyotome (Biressi et al. 2007; Kahane and Kalcheim 1998; Kahane et al. 1998). The myotome contains committed muscle cells expressing *Myf5* and *MyoD* (Kiefer and Hauschka 2001; Sassoon et al. 1989; Cinnamon et al. 2001; Ordahl et al. 2001). These two members of the basic helix-loop-helix transcription factors are considered markers of terminal specification to the muscle lineage (Pownall et al. 2002).

Only a fraction of myogenic progenitors terminally differentiate during primary myotome formation. Skeletal muscle is established in successive distinct, but overlapping, steps involving different types of myoblasts (embryonic myoblasts, fetal myoblasts, and satellite cells) (Biressi et al. 2007). The continued growth of muscles that occurs during late embryonic (E10.5–12.5 in the mouse), fetal (E14.5–17.5 in the mouse) and postnatal life is attributed to a population of muscle progenitor cells already present at the embryonic stage. These skeletal muscle progenitor cells arise in the central part of the dermomyotme, co-express *Pax3* and *Pax7*, and can differentiate into skeletal muscle fibers during embryogenesis or possibly remain as a reserve cell population within the growing muscle mass during peri- and postnatal stages (Biressi et al. 2007; Gros et al. 2005; Kassar-Duchossoy et al. 2005; Relaix et al. 2005; Schienda et al. 2006).

At around embryonic day 11 (E11) in the mouse, embryonic myoblasts invade the myotome and fuse into myotubes. At about the same stage, myogenic progenitors, which have migrated to the limb from the dermomyotome, start to differentiate into multinucleated muscles fibers (primary fibers). This process is often called primary myogenesis (Biressi et al. 2007). Myoblast differentiation is an ordered multistep process of differentiation that includes withdrawal from the cell cycle and expression of key myogenic factors, such as myogenin and MRF4 (myogenic regulatory factor 4; also known as Myf6), which will result in the

differentiation and fusion into multinucleated myotubes (Bentzinger et al. 2012; Braun and Gautel 2011). Myotubes will eventually give rise to the vast array of muscle fibers used to construct the complex skeletal muscle architecture. Most embryonic myoblasts differentiate and fuse to form primary muscle fibers, but some continue to proliferate and become fetal myoblasts. A subsequent wave of myogenesis takes place between embryonic days E14.5 and E17.5 and involves the fusion of these fetal myoblasts to form secondary fibers which are smaller than and surround the primary fibers (Biressi et al. 2007). At the end of this phase satellite cells (SCs) can be morphologically identified as mononucleated cells lying between the basal lamina and the fiber plasma membrane. In the final period of embryogenesis, muscle progenitors proliferate extensively until they reach a state in which the number of myonuclei is maintained and the synthesis of myofibrillar protein hits its peak (Schultz 1996; Sambasivan and Tajbakhsh 2007). However, during peri- and postnatal development, satellite SCs divide at a slow rate and a large part of the progeny fuse with the adjacent fiber to contribute new nuclei to growing muscle fibers (whose nuclei cannot divide). At the end of the postnatal growth period, the muscle reaches a mature state and SCs enter a phase of quiescence but can be reactivated if the muscle tissue is damaged or in response to further growth demands (Biressi et al. 2007; Bentzinger et al. 2012).

In neonatal muscle, growth is thus mainly achieved by addition of myoblasts, derived from SCs, to existing myofibers. However, while SCs are quiescent in the adult muscle, they become activated in response to injury and are able to self-renew, proliferate, and differentiate to fuse to damaged fibers or form new myofibers. Shortly after an injurious insult, inflammatory cells, such as neutrophils and macrophages, are recruited to the damaged site where they release growth factors and cytokines to induce local SCs to proliferate. The activated SCs can then divide asymmetrically and reconstitute the quiescent pool of SCs, given their capacity for self-renewal, or enter the myogenic

pathway and differentiate to restore muscle integrity at the site of injury (Yin et al. 2013). In many aspects, this *de novo* myofiber formation recapitulates embryonic myogenesis. Adult SCs, also referred to as myogenic precursor cells or adult myoblasts, express the myogenic factors *MyoD* and *Myf5*. Following proliferation, these SCs begin differentiation by down-regulating *Pax7*. The initiation of terminal differentiation and fusion begins with the expression of myogenin, which together with MyoD, activates the muscle specific structural and contractile genes. During regeneration, activated SCs have the capability to return to quiescence to maintain the SC pool, which is critical for long-term muscle integrity (Yin et al. 2013).

4.1 Panx1 Channels Mediate the Acquisition of Myogenic Commitment

As stated earlier, muscle formation is a progressive and highly orchestrated process initiated by the commitment of skeletal muscle precursor cells (Bentzinger et al. 2012). The role of Panx1 in myogenic commitment has been recently assessed using C_2C_{12} reserve cells. After differentiation of C_2C_{12} mouse myoblast cells *in vitro*, the cultures contain fused myoblasts and undifferentiated cells that have undetectable or very low levels of MyoD and Myf5 (Yoshida et al. 1998). When these undifferentiated cells are isolated and returned to growth conditions, they progress through the cell cycle and regain MyoD expression (Yoshida et al. 1998; Stuelsatz et al. 2010). As these cells once again produce both MyoD-positive differentiated and MyoD-negative undifferentiated populations when deprived of serum to induce differentiation, they have been referred to as 'reserve cells' (RCs) (Yoshida et al. 1998). The characteristics of quiescence, self-renewal, and generation of myotubes are shared by both reserve cells and the resident satellite cells of muscle fibers (Yoshida et al. 1998).

Commitment of skeletal muscle cells to differentiation is a calcium-dependent process

(Friday et al. 2000; Friday and Pavlath 2001). While the expression levels of Panx1 were not assessed directly, experiments using the Panx1 channel blockers [10]Panx (Panx1 mimetic inhibitory peptide) and carbenoxolone (CBX) indicate that Panx1 channels are present and active in uncommitted C_2C_{12} RCs and that extracellular ATP activates Ca^{2+} in these cells via P2 receptors and Panx1 channels (Riquelme et al. 2015). Extracellular ATP was shown to increase MyoD levels likely reflecting the myogenic commitment of the treated RCs (Riquelme et al. 2015) since *MyoD*, as well as *Myf5*, commit cells to the myogenic program (Bentzinger et al. 2012). This ATP-induced acquisition of myogenic commitment requires activation of P2X receptors and functional Panx1 channels. Indeed, treatment with oATP, [10]Panx, and probenecid, another Panx1 channel blocker, prevented the increase of MyoD expression induced by ATP (Riquelme et al. 2015) The induction of MyoD reactivity in the nucleus, as detected by immunofluorescence using confocal microscopy, was also prevented by knocking down Panx1 levels (Riquelme et al. 2015). Riquelme et al. propose that purinergic P2Rs and Panx1 channels are part of a positive feedback system present in C_2C_{12} RCs. They further suggest that the activation of P2XRs by extracellular ATP, released through P2Rs-activated Panx1 channels at the cell surface, contribute to the acquisition of myogenic commitment.

4.2 Panx1 Channels Promote Skeletal Muscle Myoblast Differentiation and Fusion

Amongst all development stages, differentiation is a crucial process that determines muscle cell fate and final muscle formation. Myogenic differentiation can be studied using primary cultures of skeletal muscle myoblasts or established muscle cell lines such as murine C_2C_{12} and rat L6 myoblasts. These cells can be maintained as dividing cells in serum rich medium, however under mitogen deprivation conditions (low serum, such as 2 % horse serum, media) they

exit the cell cycle, become elongated and express the various muscle-specific genes. These elongated myoblasts gradually fuse together to form large syncytial myotubes (Paterson and Strohman 1972; Yaffe and Saxel 1977; Yaffe 1968).

Using human primary skeletal muscle myoblasts (HSMM), we have shown that the levels of PANX1 protein are very low or below detectable limits by western blotting in undifferentiated myoblasts but increase drastically during their differentiation (Langlois et al. 2014). Inhibition of PANX1 channels by probenecid or CBX significantly reduced the differentiation and fusion of HSMM. Furthermore, overexpression of Panx1 accelerated HSMM differentiation (increase in the % of myoblasts that express the differentiation marker myosin heavy chain) and fusion (increase in the % of myosin heavy chain-positive cells that express two or more nuclei) when transfected cells were placed in differentiation media (2 % horse serum) (Langlois et al. 2014). However, Panx1 overexpression did not reduce HSMM proliferation when growing in serum-rich media nor did probenecid affect the inhibition of cell proliferation that is expected to occur during myoblast differentiation (Langlois et al. 2014). As exit from the cell cycle is necessary for myoblasts to enter terminal differentiation, the involvement of Panx1 in promoting myoblast differentiation is likely subsequent to the cessation of cell proliferation.

Although the mechanisms by which Panx1 promotes skeletal muscle myoblast differentiation are still unknown, the inhibition of this process by probenecid and CBX suggests an involvement of its channel activity (Langlois et al. 2014). Indeed, differentiation of C_2C_{12} cells was shown to require functional P2X receptors as well as the autocrine effect of ATP (Araya et al. 2004). It has thus been proposed that P2X receptors, probably the P2X7 type, provide an important cell influx pathway to increase the free intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) required for myogenesis (Araya et al. 2004). Since the report from Araya et al. in 2004, the role of Panx1 in ATP release from many cell

types (reviewed in (Penuela et al. 2013)), including skeletal muscle myoblasts, myotubes, and adult skeletal muscle fibers (Buvinic et al. 2009; Jorquera et al. 2012; Arias-Calderon et al. 2016), has been well described. Importantly, ATP release was demonstrated not to occur in muscles from *Panx1* knockout mice (Riquelme et al. 2013). As Panx1 is now known for its involvement in the initiation and propagation of Ca^{2+} wave signalling through P2X and P2Y receptors (reviewed in (Penuela et al. 2013)), Panx1 may play a role in the ATP release and Ca^{2+} wave propagation required for skeletal muscle myoblast differentiation. While the opening of Panx1 channels can be induced by extracellular ATP or repetitive electrical stimulation in skeletal muscle cells and fibers (Riquelme et al. 2013; Riquelme et al. 2015), the mechanisms responsible for the regulation of Panx1 expression and channel activity during the multistep process of myoblast differentiation remain to be investigated.

Once differentiation is complete, Panx1 channels are still expressed in myotubes and allow ATP release induced by repetitive electrical stimulation, which is crucial for the potentiation of skeletal muscle contraction (Riquelme et al. 2013). The ATP release through Panx1 channels has been suggested to act through P2 receptors to modulate both Ca^{2+} homeostasis and muscle physiology (Buvinic et al. 2009). Furthermore, ATP release may also be involved in the regulation of adult skeletal muscle plasticity by inducing transcriptional changes related to fast-to-slow muscle fiber phenotype transition (Jorquera et al. 2012; Arias-Calderon et al. 2016).

4.3 Panx3 Channels Regulate the Proliferation, Differentiation, and Fusion of Skeletal Muscle Myoblasts

Thus far, the only study investigating the role of Panx3 in myoblast differentiation is that recently published by our group using HSMM. While the Panx3 species of about 43 kDa were detected in human, mouse, and rat skeletal muscle tissue, they were very low or below detectable levels by western blotting in undifferentiated and differentiated HSMM lysates (Langlois et al. 2014). These species were also more abundant in adult compared to fetal skeletal muscle tissue (Langlois et al. 2014). This data suggests that Panx3 may be expressed further along the differentiation process once myotubes have become myofibers. Accordingly, Pillon et al. detected *Panx3* expression by quantitative RT-PCR in quadriceps muscle tissue, while it was not detectable in myotubes (Pillon et al. 2014). How Panx3 expression is regulated during myogenesis remains unknown, but it has been reported that the activation of the Toll-like receptor 4 (TLR4)/nuclear factor-kB (NF-kB) pathway in L6 myotubes challenged with the saturated fatty acid palmitate significantly elevates the expression of *Panx3* (Pillon et al. 2014). When the ~43 kDa Panx3 species was ectopically expressed in HSMM, it inhibited their proliferation when compared to control cells (Langlois et al. 2014). Furthermore, its overexpression also promoted HSMM differentiation and fusion based on the increase of the percentage of HSMM that were positive for myosin heavy chain and the percentage of these cells that were multinucleated (Langlois et al. 2014). These results suggest that the ~43 kDa species of Panx3 may play an important role in maintaining skeletal muscle in a differentiated and non-proliferative state.

In contrast to the ~43 kDa form of human PANX3, the ~70 kDa immunoreactive species was detected at high levels by western blotting in undifferentiated HSMM (Langlois et al. 2014). However, its levels were drastically downregulated during differentiation, and its knockdown significantly reduced HSMM proliferation (Langlois et al. 2014). While this may suggest a role for this species in keeping undifferentiated skeletal muscle myoblasts in a proliferative state, these results should be taken with caution, as stated earlier, until the exact identity of this species is adequately clarified.

5   Pannexins in Skeletal Muscle Health and Disease

While these data suggest an important function for Panxs in skeletal muscle myogenesis, as summarized in Fig. 2, the $Panx1^{-/-}$ mice surprisingly do not present with an evident skeletal muscle phenotype or defect (Riquelme et al. 2013). Interestingly, the group investigating the phenotypic effects of the global deletion of *Panx3* in mouse humeri and femora found that knockout mice had relatively larger areas of muscle attachment sites when compared with wildtype mice (Caskenette et al. 2016). Nevertheless, the skeletal muscle phenotype of both $Panx1^{-/-}$ and $Panx3^{-/-}$ mice have not been systematically examined in terms of muscle growth, myofiber number and size, fiber type distribution, overall muscle functionality and performance, etc. Furthermore, a study specifically examining skeletal muscle development from the embryonic to the adult stage using $Panx1^{-/-}$ or $Panx3^{-/-}$ mice has not been

forthcoming. It is also possible that a compensatory mechanism could occur, mitigating the effect of a lack of *Panx1* or *Panx3 in vivo*. In that sense, double knockout mice devoid of both *Panx1* and *Panx3* may be better suited to uncover the impact of Panx loss on skeletal muscle formation, growth, and function. To date, reports on the phenotype or even the existence of *Panx1/ Panx3* double knockout mice have yet to come. Based on the role of Panxs in myogenic commitment and myoblast differentiation, it would also be expected that Panx1 and Panx3 channels play an important function during skeletal muscle regeneration. Such functions could be unveiled by injecting cardiotoxin into the tibialis anterior muscle from *Panx* knockout mice and examining their regeneration as compared to cardiotoxin-injected muscles from control mice as well as those injected with saline alone (Hirata et al. 2003).

Based on the roles of Panx1 and Panx3 channels in regulating skeletal muscle myoblast proliferation and differentiation, it is also
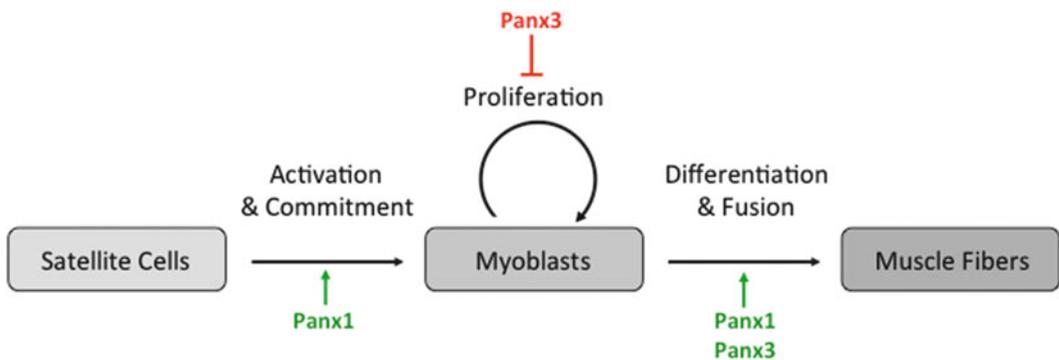
Fig. 2 **Regulation of myogenic commitment, myoblast proliferation and differentiation by Panx1 and Panx3 channels**
Activated committed satellite cells (SCs) proliferate as skeletal myoblasts before undergoing myogenic differentiation and fusion to form myofibers and eventually skeletal muscles. The acquisition of myogenic commitment by reserve cells, which are known to share many characteristics with SCs, requires functional pannexin 1 (Panx1) channels as treatment with the Panx1 channel blockers [10]Panx and probenecid prevents the increase of MyoD (myogenic transcription factor) expression (Riquelme et al. 2015). While Panx1 channels do not directly control the rate of myoblast proliferation, this process was reduced by the over-expression of the

~43 kDa species of pannexin 3 (Panx3) (Langlois et al. 2014). Interestingly, over-expression of Panx1 and Panx3 (~43 kDa) promote the differentiation and fusion of myoblast yielding myotubes *in vitro* (Langlois et al. 2014). Based on the importance of these processes in myogenesis, Panxs may be found to play critical functions during skeletal muscle development and regeneration *in vivo*. Panx1 and Panx3 may also be involved in pathologies in which skeletal muscle cell proliferation and differentiation are impaired, such as in RMS. In this cancer, tumor cells express muscle-differentiating factors but have lost the ability to terminally differentiate and thus proliferate indefinitely. However, whether Panxs levels, localization, and channel functions are altered in RMS has yet to be investigated

tempting to speculate that they may be involved in muscle pathologies in which these processes are deregulated such as that seen in rhabdomyosarcomas (RMS). RMS, an aggressive cancer with a poor prognosis, is the most common soft tissue sarcoma in children and adolescents (Loeb et al. 2008). Similar to myogenic progenitors, RMS cells express muscle-differentiating factors but have lost the ability to terminally differentiate thus proliferating indefinitely (Tapscott et al. 1993). Current therapy for RMS includes surgery, systemic chemotherapy, and ionizing radiation, but despite this aggressive approach, the prognosis of children with RMS presenting with metastasis has not improved in the last 15 years and the overall cure rate remains below 30 % (Oberlin et al. 2008). It is thought that better therapies for RMS may thus come from targeting molecular pathways that are deregulated during myogenic differentiation and as a consequence contribute to tumor formation (Wang et al. 2008; Taulli et al. 2009). As terminal myogenesis switches off cell proliferation and migration, promotion of RMS differentiation should antagonize tumor growth and metastasis (Nanni et al. 2009). In light of our recent data on the modulation of Panx1 and Panx3 levels during myogenesis and their role in inhibiting skeletal muscle myoblast proliferation and promoting differentiation, it is tempting to speculate that their levels may be deregulated in RMS. Future work should be aimed at assessing whether Panx levels are downregulated in RMS compared to healthy skeletal muscle and whether Panx channels may constitute novel therapeutic targets to treat this disease through inhibition of tumor growth and/or alleviation of its malignant properties by inducing tumor cell differentiation.

Furthermore, it has been recently found that Panx1 is part of a multiprotein complex with the dihydropyridine receptor, P2Y2 receptor, and caveolin-3 that is involved in excitation-transcription coupling of the skeletal muscle (Arias-Calderon et al. 2016). Interestingly, dystrophin also interacts with this multiprotein complex (Arias-Calderon et al. 2016). Duchenne muscular dystrophy (DMD), the most common and severe form of dystrophic muscular disease,

is caused by X-linked mutations in the dystrophin gene. Loss of dystrophin in DMD patients causes sensitivity of myofibers to mechanical damage, leading to SC activation, and myofiber regeneration. The unsustainable activation of SCs in these patients results in severe muscle wasting, infiltration of adipocytes, inflammation, and eventually paralysis and death (Liu et al. 2012; McNally and Pytel 2007). *Mdx* mice, which also have a mutation in the dystrophin gene, are commonly utilized as a model of DMD (Liu et al. 2012). Valladares et al. showed that ATP release through Panx1 channels is increased in *mdx* muscle fibers, which may be due to the higher levels of Panx1 detected in these fibers as compared to control mice (Valladares et al. 2013). Interestingly, while exogenous ATP is anti-apoptotic for normal skeletal muscle fibers, it activates pro-apoptotic pathways in fibers derived from *mdx* mice (Valladares et al. 2013). As these data suggest a role for Panx1 in the pathology of DMD and that reduction of ATP release by targeting Panx1 channels activity may be beneficial for DMD patients, such possibilities need further investigation. One approach could be the creation of a *Panx1/dystrophin* double knockout mouse with assessments looking for possible improvements in disease status as compared to control animals. By contrast, potential deregulation of Panx3 levels and/or functions in *mdx* fibers have yet to be examined.

6   Concluding Remarks

Only recently have the roles for Panxs in regulating myogenic commitment of reserve cells as well as the proliferation and differentiation of skeletal muscle myoblasts been unveiled *in vitro*. While these studies suggest important roles for Panxs in skeletal muscle health and disease, much work remains to be done in order to decipher the mechanism regulating Panx levels, post-translational modifications, and functions during the various steps of myogenesis, as well as to elucidate the molecular pathway by which Panxs regulate this process. Furthermore,

more research is needed to specifically address the potential roles of Panx1, and possibly Panx3, on satellite cell quiescence and self-renewal. Based on the importance of myogenic commitment and myoblast proliferation and differentiation in skeletal muscle development, regeneration, RMS, and dystrophy, it is expected that roles for Panx1 and Panx3 in crucial physiological and pathological processes of skeletal muscle will emerge in the years to come. Now that *Panx1* and *Panx3* null mice are available in several laboratories, future studies may reveal whether Panxs play a role in myogenesis and skeletal muscle regeneration *in vivo* and whether Panxs represent potential new therapeutic targets for skeletal muscle diseases.

Conflicts of Interest No conflicts of interest, financial or otherwise, are declared by the authors.

References

Araya R, Riquelme MA, Brandan E, Saez JC (2004) The formation of skeletal muscle myotubes requires functional membrane receptors activated by extracellular ATP. Brain Res Brain Res Rev 47(1–3):174–188. doi: S0165017304000761 [pii] 10.1016/j.brainresrev. 2004.06.003

Arias-Calderon M, Almarza G, Diaz-Vegas A, Contreras-Ferrat A, Valladares D, Casas M, Toledo H, Jaimovich E, Buvinic S (2016) Characterization of a multiprotein complex involved in excitation-transcription coupling of skeletal muscle. Skelet Muscle 6:15. doi:10.1186/s13395-016-0087-5

Bao L, Locovei S, Dahl G (2004) Pannexin membrane channels are mechanosensitive conduits for ATP. FEBS Lett 572(1–3):65–68

Baranova A, Ivanov D, Petrash N, Pestova A, Skoblov M, Kelmanson I, Shagin D, Nazarenko S, Geraymovych E, Litvin O, Tiunova A, Born TL, Usman N, Staroverov D, Lukyanov S, Panchin Y (2004) The mammalian pannexin family is homologous to the invertebrate innexin gap junction proteins. Genomics 83(4):706–716

Bargiotas P, Krenz A, Hormuzdi SG, Ridder DA, Herb A, Barakat W, Penuela S, von Engelhardt J, Monyer H, Schwaninger M (2011) Pannexins in ischemia-induced neurodegeneration. Proc Natl Acad Sci U S A 108(51):20772–20777. doi:10.1073/pnas. 1018262108

Beckmann A, Grissmer A, Krause E, Tschernig T, Meier C (2016) Pannexin-1 channels show distinct morphology and no gap junction characteristics in mammalian cells. Cell Tissue Res 363(3):751–763. doi:10.1007/s00441-015-2281-x

Bentzinger CF, Wang YX, Rudnicki MA (2012) Building muscle: molecular regulation of myogenesis. Cold Spring Harb Perspect Biol 4(2):pii: a008342. doi:10.1101/cshperspect.a008342

Biressi S, Molinaro M, Cossu G (2007) Cellular heterogeneity during vertebrate skeletal muscle development. Dev Biol 308(2):281–293. doi:10.1016/j.ydbio. 2007.06.006

Boassa D, Ambrosi C, Qiu F, Dahl G, Gaietta G, Sosinsky G (2007) Pannexin1 channels contain a glycosylation site that targets the hexamer to the plasma membrane. J Biol Chem 282(43):31733–31743

Bond SR, Wang N, Leybaert L, Naus CC (2012) Pannexin 1 ohnologs in the teleost lineage. J Membr Biol 245 (8):483–493. doi:10.1007/s00232-012-9497-4

Braun T, Gautel M (2011) Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. Nat Rev Mol Cell Biol 12(6):349–361. doi:nrm3118 [pii] 10.1038/nrm3118

Bruzzone R, Hormuzdi SG, Barbe MT, Herb A, Monyer H (2003) Pannexins, a family of gap junction proteins expressed in brain. Proc Natl Acad Sci U S A 100 (23):13644–13649

Buvinic S, Almarza G, Bustamante M, Casas M, Lopez J, Riquelme M, Saez JC, Huidobro-Toro JP, Jaimovich E (2009) ATP released by electrical stimuli elicits calcium transients and gene expression in skeletal muscle. J Biol Chem 284(50):34490–34505. doi: M109.057315 [pii] 10.1074/jbc.M109.057315

Caskenette D, Penuela S, Lee V, Barr K, Beier F, Laird DW, Willmore KE (2016) Global deletion of Panx3 produces multiple phenotypic effects in mouse humeri and femora. J Anat. doi:10.1111/joa.12437

Cea LA, Cisterna BA, Puebla C, Frank M, Figueroa XF, Cardozo C, Willecke K, Latorre R, Saez JC (2013) De novo expression of connexin hemichannels in denervated fast skeletal muscles leads to atrophy. Proc Natl Acad Sci U S A 110(40):16229–16234. doi:10.1073/pnas.1312331110

Celetti SJ, Cowan KN, Penuela S, Shao Q, Churko J, Laird DW (2010) Implications of pannexin 1 and pannexin 3 for keratinocyte differentiation. J Cell Sci 123(Pt 8):1363–1372. doi:jcs.056093 [pii] 10.1242/jcs.056093

Chekeni FB, Elliott MR, Sandilos JK, Walk SF, Kinchen JM, Lazarowski ER, Armstrong AJ, Penuela S, Laird DW, Salvesen GS, Isakson BE, Bayliss DA, Ravichandran KS (2010) Pannexin 1 channels mediate 'find-me' signal release and membrane permeability during apoptosis. Nature 467(7317):863–867. doi: nature09413 [pii] 10.1038/nature09413

Christ B, Ordahl CP (1995) Early stages of chick somite development. Anat Embryol (Berl) 191(5):381–396

Cinnamon Y, Kahane N, Bachelet I, Kalcheim C (2001) The sub-lip domain--a distinct pathway for myotome precursors that demonstrate rostral-caudal migration. Development 128(3):341–351

Cowan KN, Langlois S, Penuela S, Cowan BJ, Laird DW (2012) Pannexin1 and Pannexin3 exhibit distinct localization patterns in human skin appendages and are regulated during keratinocyte differentiation and carcinogenesis. Cell Commun Adhes 19(3–4):45–53. doi:10.3109/15419061.2012.712575

D'Hondt C, Ponsaerts R, De Smedt H, Vinken M, De Vuyst E, De Bock M, Wang N, Rogiers V, Leybaert L, Himpens B, Bultynck G (2011) Pannexin channels in ATP release and beyond: an unexpected rendezvous at the endoplasmic reticulum. Cell Signal 23(2):305–316. doi:S0898-6568(10)00211-1 [pii] 10.1016/j.cellsig.2010.07.018

Diezmos EF, Sandow SL, Markus I, Shevy Perera D, Lubowski DZ, King DW, Bertrand PP, Liu L (2013) Expression and localization of pannexin-1 hemichannels in human colon in health and disease. Neurogastroenterol Motil 25(6):e395–e405. doi:10.1111/nmo.12130

Dolmatova E, Spagnol G, Boassa D, Baum JR, Keith K, Ambrosi C, Kontaridis MI, Sorgen PL, Sosinsky GE, Duffy HS (2012) Cardiomyocyte ATP release through pannexin 1 aids in early fibroblast activation. Am J Physiol Heart Circ Physiol 303(10):H1208–H1218. doi:10.1152/ajpheart.00251.2012

Ednie AR, Bennett ES (2012) Modulation of voltage-gated ion channels by sialylation. Compr Physiol 2 (2):1269–1301. doi:10.1002/cphy.c110044

Friday BB, Pavlath GK (2001) A calcineurin- and NFAT-dependent pathway regulates Myf5 gene expression in skeletal muscle reserve cells. J Cell Sci 114 (Pt 2):303–310

Friday BB, Horsley V, Pavlath GK (2000) Calcineurin activity is required for the initiation of skeletal muscle differentiation. J Cell Biol 149(3):657–666

Goulding MD, Chalepakis G, Deutsch U, Erselius JR, Gruss P (1991) Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. Embo J 10(5):1135–1147

Gros J, Manceau M, Thome V, Marcelle C (2005) A common somitic origin for embryonic muscle progenitors and satellite cells. Nature 435 (7044):954–958. doi:10.1038/nature03572

Gulbransen BD, Sharkey KA (2012) Novel functional roles for enteric glia in the gastrointestinal tract. Nat Rev Gastroenterol Hepatol 9(11):625–632. doi:nrgastro.2012.138 [pii] 10.1038/nrgastro.2012.138

Hirata A, Masuda S, Tamura T, Kai K, Ojima K, Fukase A, Motoyoshi K, Kamakura K, Miyagoe-Suzuki Y, Takeda S (2003) Expression profiling of cytokines and related genes in regenerating skeletal muscle after cardiotoxin injection: a role for osteopontin. Am J Pathol 163(1):203–215. doi:10.1016/S0002-9440(10)63644-9

Huang YJ, Maruyama Y, Dvoryanchikov G, Pereira E, Chaudhari N, Roper SD (2007) The role of pannexin 1 hemichannels in ATP release and cell-cell communication in mouse taste buds. Proc Natl Acad Sci U S A 104(15):6436–6441

Ishikawa M, Iwamoto T, Nakamura T, Doyle A, Fukumoto S, Yamada Y (2011) Pannexin 3 functions as an ER Ca2+ channel, hemichannel, and gap junction to promote osteoblast differentiation. J Cell Biol 193(7):1257–1274. doi:jcb.201101050 [pii] 10.1083/jcb.201101050

Ishikawa M, Iwamoto T, Fukumoto S, Yamada Y (2014) Pannexin 3 Inhibits Proliferation of Osteoprogenitor Cells by Regulating Wnt and p21 Signaling. J Biol Chem 289(5):2839–2851. doi:M113.523241 [pii] 10.1074/jbc.M113.523241

Ishikawa M, Williams GL, Ikeuchi T, Sakai K, Fukumoto S, Yamada Y (2016) Pannexin 3 and connexin 43 modulate skeletal development through their distinct functions and expression patterns. J Cell Sci 129(5):1018–1030. doi:10.1242/jcs.176883

Iwamoto T, Nakamura T, Doyle A, Ishikawa M, de Vega S, Fukumoto S, Yamada Y (2010) Pannexin 3 regulates intracellular ATP/cAMP levels and promotes chondrocyte differentiation. J Biol Chem 285(24):18948–18958. doi:M110.127027 [pii] 10.1074/jbc.M110.127027

Jorquera G, Altamirano F, Contreras-Ferrat A, Almarza G, Buvinic S, Jacquemond V, Jaimovich E, Casas M (2012) Cav1.1 controls frequency-dependent events regulating adult skeletal muscle plasticity. J Cell Sci 126(Pt 5):1189–1198. doi:jcs.116855 [pii] 10.1242/jcs.116855

Jostes B, Walther C, Gruss P (1990) The murine paired box gene, Pax7, is expressed specifically during the development of the nervous and muscular system. Mech Dev 33(1):27–37

Kahane N, Kalcheim C (1998) Identification of early postmitotic cells in distinct embryonic sites and their possible roles in morphogenesis. Cell Tissue Res 294 (2):297–307

Kahane N, Cinnamon Y, Kalcheim C (1998) The cellular mechanism by which the dermomyotome contributes to the second wave of myotome development. Development 125(21):4259–4271

Kassar-Duchossoy L, Giacone E, Gayraud-Morel B, Jory A, Gomes D, Tajbakhsh S (2005) Pax3/Pax7 mark a novel population of primitive myogenic cells during development. Genes Dev 19(12):1426–1431. doi:10.1101/gad.345505

Kiefer JC, Hauschka SD (2001) Myf-5 is transiently expressed in nonmuscle mesoderm and exhibits dynamic regional changes within the presegmented mesoderm and somites I-IV. Dev Biol 232(1):77–90. doi:10.1006/dbio.2000.0114

Lai CP, Bechberger JF, Thompson RJ, MacVicar BA, Bruzzone R, Naus CC (2007) Tumor-suppressive effects of pannexin 1 in C6 glioma cells. Cancer Res 67(4):1545–1554

Lai CP, Bechberger JF, Naus CC (2009) Pannexin2 as a novel growth regulator in C6 glioma cells. Oncogene 28(49):4402–4408

Langlois S, Xiang X, Young K, Cowan BJ, Penuela S, Cowan KN (2014) Pannexin 1 and pannexin

3 channels regulate skeletal muscle myoblast proliferation and differentiation. J Biol Chem 289 (44):30717–30731. doi:M114.572131 [pii] 10.1074/jbc.M114.572131

Le Vasseur M, Lelowski J, Bechberger JF, Sin WC, Naus CC (2014) Pannexin 2 protein expression is not restricted to the CNS. Front Cell Neurosci 8:392. doi:10.3389/fncel.2014.00392

Liu N, Williams AH, Maxeiner JM, Bezprozvannaya S, Shelton JM, Richardson JA, Bassel-Duby R, Olson EN (2012) microRNA-206 promotes skeletal muscle regeneration and delays progression of Duchenne muscular dystrophy in mice. J Clin Invest 122 (6):2054–2065. doi:62656 [pii] 10.1172/JCI62656

Locovei S, Bao L, Dahl G (2006a) Pannexin 1 in erythrocytes: function without a gap. Proc Natl Acad Sci U S A 103(20):7655–7659

Locovei S, Wang J, Dahl G (2006b) Activation of pannexin 1 channels by ATP through P2Y receptors and by cytoplasmic calcium. FEBS Lett 580 (1):239–244

Loeb DM, Thornton K, Shokek O (2008) Pediatric soft tissue sarcomas. Surg Clin North Am 88(3):615–627, vii. doi:S0039-6109(08)00042-X [pii] 10.1016/j.suc.2008.03.008

Lohman AW, Weaver JL, Billaud M, Sandilos JK, Griffiths R, Straub AC, Penuela S, Leitinger N, Laird DW, Bayliss DA, Isakson BE (2012) S-nitrosylation inhibits pannexin 1 channel function. J Biol Chem 287 (47):39602–39612. doi:M112.397976 [pii] 10.1074/jbc.M112.397976

Lohman AW, Leskov IL, Butcher JT, Johnstone SR, Stokes TA, Begandt D, DeLalio LJ, Best AK, Penuela S, Leitinger N, Ravichandran KS, Stokes KY, Isakson BE (2015) Pannexin 1 channels regulate leukocyte emigration through the venous endothelium during acute inflammation. Nat Commun 6:7965. doi:10.1038/ncomms8965

McNally EM, Pytel P (2007) Muscle diseases: the muscular dystrophies. Annu Rev Pathol 2:87–109. doi:10.1146/annurev.pathol.2.010506.091936

Moon PM, Penuela S, Barr K, Khan S, Pin CL, Welch I, Attur M, Abramson SB, Laird DW, Beier F (2015) Deletion of Panx3 Prevents the Development of Surgically Induced Osteoarthritis. J Mol Med (Berl). doi:10.1007/s00109-015-1311-1

Nanni P, Nicoletti G, Palladini A, Astolfi A, Rinella P, Croci S, Landuzzi L, Monduzzi G, Stivani V, Antognoli A, Murgo A, Ianzano M, De Giovanni C, Lollini PL (2009) Opposing control of rhabdomyosarcoma growth and differentiation by myogenin and interleukin 4. Mol Cancer Ther 8(4):754–761. doi:1535-7163.MCT-08-0678 [pii] 10.1158/1535-7163.MCT-08-0678

Oberlin O, Rey A, Lyden E, Bisogno G, Stevens MC, Meyer WH, Carli M, Anderson JR (2008) Prognostic factors in metastatic rhabdomyosarcomas: results of a pooled analysis from United States and European cooperative groups. J Clin Oncol 26(14):2384–2389. doi:26/14/2384 [pii] 10.1200/JCO.2007.14.7207

Oh SK, Shin JO, Baek JI, Lee J, Bae JW, Ankamerddy H, Kim MJ, Huh TL, Ryoo ZY, Kim UK, Bok J, Lee KY (2015) Pannexin 3 is required for normal progression of skeletal development in vertebrates.. doi:fj.15-273722 [pii] 10.1096/fj.15-273722

Ordahl CP, Berdougo E, Venters SJ, Denetclaw WF Jr (2001) The dermomyotome dorsomedial lip drives growth and morphogenesis of both the primary myotome and dermomyotome epithelium. Development 128(10):1731–1744

Orellana JA, Velasquez S, Williams DW, Saez JC, Berman JW, Eugenin EA (2013) Pannexin1 hemichannels are critical for HIV infection of human primary CD4+ T lymphocytes. J Leukoc Biol 94 (3):399–407. doi:10.1189/jlb.0512249

Panchin YV (2005) Evolution of gap junction proteins-- the pannexin alternative. J Exp Biol 208 (Pt 8):1415–1419

Panchin Y, Kelmanson I, Matz M, Lukyanov K, Usman N, Lukyanov S (2000) A ubiquitous family of putative gap junction molecules. Curr Biol 10(13): R473–R474. doi:S0960-9822(00)00576-5 [pii]

Paoletti A, Raza SQ, Voisin L, Law F, Caillet M, Martins I, Deutsch E, Perfettini JL (2013) Editorial: Pannexin-1--the hidden gatekeeper for HIV-1. J Leukoc Biol 94(3):390–392. doi:10.1189/jlb.0313148

Paterson B, Strohman RC (1972) Myosin synthesis in cultures of differentiating chicken embryo skeletal muscle. Dev Biol 29(2):113–138

Pelegrin P, Surprenant A (2006) Pannexin-1 mediates large pore formation and interleukin-1beta release by the ATP-gated P2X7 receptor. Embo J 25 (21):5071–5082

Penuela S, Bhalla R, Gong XQ, Cowan KN, Celetti SJ, Cowan BJ, Bai D, Shao Q, Laird DW (2007) Pannexin 1 and pannexin 3 are glycoproteins that exhibit many distinct characteristics from the connexin family of gap junction proteins. J Cell Sci 120 (Pt 21):3772–3783. doi:jcs.009514 [pii] 10.1242/jcs.009514

Penuela S, Bhalla R, Nag K, Laird DW (2009) Glycosylation Regulates Pannexin Intermixing and Cellular Localization. Mol Biol Cell 20(20):4313–4323

Penuela S, Gehi R, Laird DW (2013) The biochemistry and function of pannexin channels. Biochim Biophys Acta 1828(1):15–22. doi:S0005-2736(12)00021-1 [pii] 10.1016/j.bbamem.2012.01.017

Penuela S, Kelly JJ, Churko JM, Barr KJ, Berger AC, Laird DW (2014a) Panx1 regulates cellular properties of keratinocytes and dermal fibroblasts in skin development and wound healing. J Invest Dermatol 134 (7):2026–2035. doi:10.1038/jid.2014.86

Penuela S, Lohman AW, Lai W, Gyenis L, Litchfield DW, Isakson BE, Laird DW (2014b) Diverse post-translational modifications of the pannexin family of

channel-forming proteins. Channels (Austin) 8 (2):124–130. doi:27422 [pii] 10.4161/chan.27422

Pillon NJ, Li YE, Fink LN, Brozinick JT, Nikolayev A, Kuo MS, Bilan PJ, Klip A (2014) Nucleotides released from palmitate-challenged muscle cells through pannexin-3 attract monocytes. Diabetes 63 (11):3815–3826. doi:db14-0150 [pii] 10.2337/db14-0150

Pinheiro AR, Paramos-de-Carvalho D, Certal M, Costa MA, Costa C, Magalhaes-Cardoso MT, Ferreirinha F, Sevigny J, Correia-de-Sa P (2013) Histamine induces ATP release from human subcutaneous fibroblasts, via pannexin-1 hemichannels, leading to Ca2+ mobilization and cell proliferation. J Biol Chem 288 (38):27571–27583. doi:10.1074/jbc.M113.460865

Pownall ME, Gustafsson MK, Emerson CP Jr (2002) Myogenic regulatory factors and the specification of muscle progenitors in vertebrate embryos. Annu Rev Cell Dev Biol 18:747–783. doi:10.1146/annurev.cellbio.18.012502.105758

Prochnow N, Abdulazim A, Kurtenbach S, Wildforster V, Dvoriantchikova G, Hanske J, Petrasch-Parwez E, Shestopalov VI, Dermietzel R, Manahan-Vaughan D, Zoidl G (2012) Pannexin1 stabilizes synaptic plasticity and is needed for learning. PLoS ONE 7(12): e51767. doi:10.1371/journal.pone.0051767 PONE-D-12-23109 [pii]

Qu Y, Misaghi S, Newton K, Gilmour LL, Louie S, Cupp JE, Dubyak GR, Hackos D, Dixit VM (2011) Pannexin-1 is required for ATP release during apoptosis but not for inflammasome activation. J Immunol 186(11):6553–6561. doi:jimmunol.1100478 [pii] 10.4049/jimmunol.1100478

Relaix F, Rocancourt D, Mansouri A, Buckingham M (2005) A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. Nature 435(7044):948–953. doi:10.1038/nature03594

Riquelme MA, Cea LA, Vega JL, Boric MP, Monyer H, Bennett MV, Frank M, Willecke K, Saez JC (2013) The ATP required for potentiation of skeletal muscle contraction is released via pannexin hemichannels. Neuropharmacology 75:594–603. doi:S0028-3908 (13)00118-4 [pii] 10.1016/j.neuropharm.2013.03.022

Riquelme MA, Cea LA, Vega JL, Puebla C, Vargas AA, Shoji KF, Subiabre M, Saez JC (2015) Pannexin channels mediate the acquisition of myogenic commitment in C2C12 reserve cells promoted by P2 receptor activation. Front Cell Dev Biol 3:25. doi:10.3389/fcell.2015.00025

Sambasivan R, Tajbakhsh S (2007) Skeletal muscle stem cell birth and properties. Semin Cell Dev Biol 18 (6):870–882. doi:10.1016/j.semcdb.2007.09.013

Sassoon D, Lyons G, Wright WE, Lin V, Lassar A, Weintraub H, Buckingham M (1989) Expression of two myogenic regulatory factors myogenin and MyoD1 during mouse embryogenesis. Nature 341 (6240):303–307. doi:10.1038/341303a0

Scaal M, Christ B (2004) Formation and differentiation of the avian dermomyotome. Anat Embryol (Berl) 208 (6):411–424. doi:10.1007/s00429-004-0417-y

Schienda J, Engleka KA, Jun S, Hansen MS, Epstein JA, Tabin CJ, Kunkel LM, Kardon G (2006) Somitic origin of limb muscle satellite and side population cells. Proc Natl Acad Sci U S A 103(4):945–950. doi:10.1073/pnas.0510164103

Schultz E (1996) Satellite cell proliferative compartments in growing skeletal muscles. Dev Biol 175(1):84–94. doi:10.1006/dbio.1996.0097

Schwetz TA, Norring SA, Ednie AR, Bennett ES (2011) Sialic acids attached to O-glycans modulate voltage-gated potassium channel gating. J Biol Chem 286 (6):4123–4132. doi:10.1074/jbc.M110.171322

Seror C, Melki MT, Subra F, Raza SQ, Bras M, Saidi H, Nardacci R, Voisin L, Paoletti A, Law F, Martins I, Amendola A, Abdul-Sater AA, Ciccosanti F, Delelis O, Niedergang F, Thierry S, Said-Sadier N, Lamaze C, Metivier D, Estaquier J, Fimia GM, Falasca L, Casetti R, Modjtahedi N, Kanellopoulos J, Mouscadet JF, Ojcius DM, Piacentini M, Gougeon ML, Kroemer G, Perfettini JL (2011) Extracellular ATP acts on P2Y2 purinergic receptors to facilitate HIV-1 infection. J Exp Med 208(9):1823–1834. doi:10.1084/jem.20101805

Shestopalov VI, Panchin Y (2008) Pannexins and gap junction protein diversity. Cell Mol Life Sci 65 (3):376–394

Silverman WR, de Rivero Vaccari JP, Locovei S, Qiu F, Carlsson SK, Scemes E, Keane RW, Dahl G (2009) The pannexin 1 channel activates the inflammasome in neurons and astrocytes. J Biol Chem 284 (27):18143–18151

Sosinsky GE, Boassa D, Dermietzel R, Duffy HS, Laird DW, MacVicar B, Naus CC, Penuela S, Scemes E, Spray DC, Thompson RJ, Zhao HB, Dahl G (2011) Pannexin channels are not gap junction hemichannels. Channels (Austin) 5(3):193–197. doi:15765 [pii]

Stewart MK, Plante I, Penuela S, Laird DW (2016) Loss of Panx1 Impairs Mammary Gland Development at Lactation: Implications for Breast Tumorigenesis. PLoS ONE 11(4):e0154162. doi:10.1371/journal.pone.0154162

Stuelsatz P, Pouzoulet F, Lamarre Y, Dargelos E, Poussard S, Leibovitch S, Cottin P, Veschambre P (2010) Down-regulation of MyoD by calpain 3 promotes generation of reserve cells in C2C12 myoblasts. J Biol Chem 285(17):12670–12683. doi:10.1074/jbc.M109.063966

Suadicani SO, Iglesias R, Wang J, Dahl G, Spray DC, Scemes E (2012) ATP signaling is deficient in cultured Pannexin1-null mouse astrocytes. Glia 60 (7):1106–1116. doi:10.1002/glia.22338

Swayne LA, Sorbara CD, Bennett SA (2010) Pannexin 2 is expressed by postnatal hippocampal neural progenitors and modulates neuronal commitment. J Biol Chem 285(32):24977–24986. doi:M110.130054 [pii] 10.1074/jbc.M110.130054

Tapscott SJ, Thayer MJ, Weintraub H (1993) Deficiency in rhabdomyosarcomas of a factor required for MyoD activity and myogenesis. Science 259 (5100):1450–1453

Taulli R, Bersani F, Foglizzo V, Linari A, Vigna E, Ladanyi M, Tuschl T, Ponzetto C (2009) The muscle-specific microRNA miR-206 blocks human rhabdomyosarcoma growth in xenotransplanted mice by promoting myogenic differentiation. J Clin Invest 119(8):2366–2378. doi:38075 [pii] 10.1172/JCI38075

Thompson RJ, Zhou N, MacVicar BA (2006) Ischemia opens neuronal gap junction hemichannels. Science 312(5775):924–927

Thompson RJ, Jackson MF, Olah ME, Rungta RL, Hines DJ, Beazely MA, MacDonald JF, MacVicar BA (2008) Activation of pannexin-1 hemichannels augments aberrant bursting in the hippocampus. Science 322(5907):1555–1559

Timoteo MA, Carneiro I, Silva I, Noronha-Matos JB, Ferreirinha F, Silva-Ramos M, Correia-de-Sa P (2014) ATP released via pannexin-1 hemichannels mediates bladder overactivity triggered by urothelial P2Y6 receptors. Biochem Pharmacol 87(2):371–379. doi:10.1016/j.bcp.2013.11.007

Turmel P, Dufresne J, Hermo L, Smith CE, Penuela S, Laird DW, Cyr DG (2011) Characterization of pannexin1 and pannexin3 and their regulation by androgens in the male reproductive tract of the adult rat. Mol Reprod Dev 78(2):124–138. doi:10.1002/mrd.21280

Valladares D, Almarza G, Contreras A, Pavez M, Buvinic S, Jaimovich E, Casas M (2013) Electrical stimuli are anti-apoptotic in skeletal muscle via extracellular ATP. Alteration of this signal in Mdx mice is a likely cause of dystrophy. PLoS ONE 8(11):e75340. doi:10.1371/journal.pone.0075340

Vanden Abeele F, Bidaux G, Gordienko D, Beck B, Panchin YV, Baranova AV, Ivanov DV, Skryma R, Prevarskaya N (2006) Functional implications of calcium permeability of the channel formed by pannexin 1. J Cell Biol 174(4):535–546

Wang H, Garzon R, Sun H, Ladner KJ, Singh R, Dahlman J, Cheng A, Hall BM, Qualman SJ, Chandler DS, Croce CM, Guttridge DC (2008) NF-kappaB-YY1-miR-29 regulatory circuitry in skeletal myogenesis and rhabdomyosarcoma. Cancer Cell 14(5):369–381. doi:S1535-6108(08)00330-9 [pii] 10.1016/j.ccr.2008.10.006

Weilinger NL, Lohman AW, Rakai BD, Ma EM, Bialecki J, Maslieieva V, Rilea T, Bandet MV, Ikuta NT, Scott L, Colicos MA, Teskey GC, Winship IR, Thompson RJ (2016) Metabotropic NMDA receptor signaling couples Src family kinases to pannexin-1 during excitotoxicity. Nat Neurosci 19(3):432–442. doi:10.1038/nn.4236

Yaffe D (1968) Retention of differentiation potentialities during prolonged cultivation of myogenic cells. Proc Natl Acad Sci U S A 61(2):477–483

Yaffe D, Saxel O (1977) Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. Nature 270(5639):725–727

Yin H, Price F, Rudnicki MA (2013) Satellite cells and the muscle stem cell niche. Physiol Rev 93(1):23–67. doi:10.1152/physrev.00043.2011

Yoshida N, Yoshida S, Koishi K, Masuda K, Nabeshima Y (1998) Cell heterogeneity upon myogenic differentiation: down-regulation of MyoD and Myf-5 generates 'reserve cells'. J Cell Sci 111(Pt 6):769–779

Zoidl G, Kremer M, Zoidl C, Bunse S, Dermietzel R (2008) Molecular diversity of connexin and pannexin genes in the retina of the zebrafish Danio rerio. Cell Commun Adhes 15(1):169–183

Hyaluronidase and Chondroitinase

Wenshuang Wang, Junhong Wang, and Fuchuan Li

Abstract

Glycosaminoglycans (GAGs) are important constituents of the extracellular matrix that make significant contributions to biological processes and have been implicated in a wide variety of diseases. GAG-degrading enzymes with different activities have been found in various animals and microorganisms, and they play an irreplaceable role in the structure and function studies of GAGs. As two kind of important GAG-degrading enzymes, hyaluronidase (HAase) and chondroitinase (CSase) have been widely studied and increasing evidence has shown that, in most cases, their substrate specificities overlap and thus the "HAase" or "CSase" terms may be improper or even misnomers. Different from previous reviews, this article combines HAase and CSase together to discuss the traditional classification, substrate specificity, degradation pattern, new resources and naming of these enzymes.

Keyword

Glycosaminoglycan • Hyaluronic acid • Chondroitin sulfate/dermatan sulfate • Hyaluronidase • Chondroitinase

Abbreviations

GAG	glycosaminoglycan
CS	chondroitin sulfate
DS	dermatan sulfate

HA	hyaluronic acid
Hep	heparin
HS	heparan sulfate
GlcUA	D-glucuronic acid
IdoUA	L-iduronic acid
HexUA	hexuronic acid
CSase	chondroitinase
HAase	hyaluronidase

W. Wang and F. Li (✉)
National Glycoengineering Research Center, Shandong University, 27 S. Shanda Road, Jinan 250100, China

Shenzhen Research Institute of Shandong University, Rm A301, Virtual University Park, Nanshan, Shenzhen 518057, China
e-mail: fuchuanli@sdu.edu.cn

J. Wang
School of Foreign Languages, University of Shanghai for Science and Technology, 516 Jungong Road, Shanghai 200093, China

GalNAc   N-acetyl-D-galactosamine
GlcNAc   N-acetyl-D-glucosamine
HexUA   hexuronic acid
HexN   hexosamine

1 Glycosaminoglycans

Glycosaminoglycans (GAGs) are a family of negatively charged and linear heteropolysaccharides that are ubiquitously distributed on cell surfaces and in extracellular matrices of animals. GAGs are involved in a wide array of biological processes (Boneu 1995; Jackson et al. 1991), including cell adhesion (Sugahara et al. 2003; Sugahara and Mikami 2007; Handel et al. 2005; Bülow and Hobert 2006), neurite outgrowth promotion (Faissner et al. 1994; Clement et al. 1998), cell proliferation (Sugahara et al. 2003; Bülow and Hobert 2006), tissue morphogenesis (Klüppel et al. 2005), viral infection (Hsiao et al. 1999; Williams and Straus 1997; Bergefall et al. 2005), regulation of the signaling of various growth factors (such as aFGF, bFGF, HGF, PTN and so on) (Nandi et al. 2006; Taylor and Gallo 2006) and cytokines (Hwang et al. 2003; Mizuguchi et al. 2003; Izumikawa et al. 2004; Mizumoto and Sugahara 2013). Based on their chemical structure, GAGs are classified into four basic families: hyaluronic acid (HA), chondroitin sulfate/dermatan sulfate (CS/DS), heparin/heparan sulfate (Hep/HS) and keratan sulfate (Capila and Linhardt 2002; Esko and Selleck 2002; Linhardt and Toida 2004) (Fig. 1). This review mainly focuses on HA/CS/DS-degrading enzymes.

HA was first isolated and identified from cattle eyes in 1934 (Meyer and Palmer 1934) and was subsequently shown to exist widely in connective, epithelial and neural tissue (Kamhi et al. 2013). Different from other GAGs, HA is the simplest GAG that neither contains sulfate groups nor attaches covalently to a core protein to form proteoglycans. The molecular weight of HA in nature is usually much larger than other GAGs (100–1000 kDa), and it is composed of the repeating disaccharide unit GlcUAβ1-3GlcNAc, in which GlcUA and GlcNAc represent D-glucuronic acid and N-acetyl-D-glucosamine, respectively (Fraser et al. 1997).
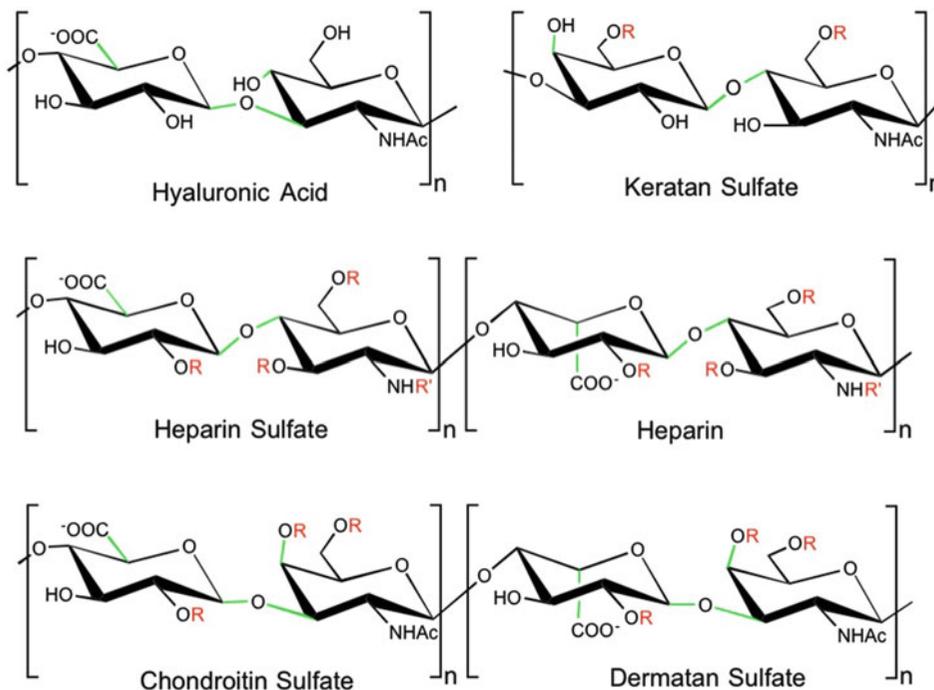

Fig. 1 The chemical structure of GAG. (R: H or SO_3^-, R': H or Ac)

The CS/DS chain is composed of repeating disaccharide units consisting of GlcUA or L-iduronic acid (IdoUA) glycosidically linked to GalNAc (GlcUAβ1-3GalNAc or IdoUAα1-3GalNAc) (Sugahara et al. 2003). In biosynthesis, during or after CS chains polymerization, CS chains are further modified by differential sulfation by specific sulfotransferases at C-2 of GlcUA/IdoUA and/or C-4 and/or C-6 of GalNAc to yield prominent structural diversity (Kusche-Gullberg and Kjellén 2003). Furthermore, some GlcUA residues are epimerized into IdoUA by the action of glucuronyl C-5 epimerase, and the chain that contains repeating disaccharide units of -IdoUA-GalNAc- has been designated as DS (Silbert and Sugumaran 2002; Maccarana et al. 2006). Thus, the two chains are often detected as CS/DS co-polymeric structures (CS-DS) and are usually periodically distributed in a cell/tissue-specific manner (Izumikawa et al. 2004; Cheng et al. 1994). Differences in the sulfation pattern result in a variety of disaccharide units including O/iO unit [GlcUAβ1-3GalNAc/IdoUAα1-3GalNAc], A/iA unit [GlcUAβ1-3GalNAc(4S)/IdoUAα1-3GalNAc(4S)], B/iB unit [GlcUA(2S)β1-3GalNAc(4S)/IdoUA(2S)α1-3GalNAc(4S)], C/iC unit [GlcUAβ1-3GalNAc(6S)/IdoUAα1-3GalNAc(6S)], D/iD unit [GlcUA(2S)β1-3GalNAc(6S)/IdoUA(2S)α1-3GalNAc(6S)], E/iE unit [GlcUAβ1-3GalNAc(4S,6S)/IdoUAα1-3GalNAc(4S,6S)] and T/iT unit [GlcUA(2S)β1-3GalNAc(4S,6S)/IdoUA(2S)α1-3GalNAc(4S,6S)], where 2S, 4S and 6S represent 2-O-, 4-O- and 6-O-sulfate groups, respectively (Sugahara et al. 2003; Sugahara and Mikami 2007).

The CS/DS chains have the potential to display an enormous microheterogeneity, which is the structural basis for various biological functions. However, the high complexity of the structure brings enormous difficulty to structural and functional studies of CS/DS chains. As in the case of other GAGs, CS/DS-degrading enzymes with different specific activities are powerful tools to solve this problem. Such enzymes can be used alone or combined with spectroscopic techniques such as nuclear magnetic resonance

(NMR) and mass spectrometry (MS) in the structural and functional analyses of CS/DS chains (Linhardt et al. 2006; Li et al. 2010; Nadanaka et al. 1998).

2 CS/DS and HA-Degrading Enzymes

Based on the enzymatic mechanism, glycosaminoglycan degradation enzymes can be divided into two categories: hydrolases and lyases. Hydrolases are assigned to glycoside hydrolase (GH) families based on amino acid sequence similarities (Henrissat 1991). The hexuronic acid-hexosamine bond of GAGs is cleaved by hydrolase through a standard mechanism of glycosidase, in which the glycosidic bond is hydrolyzed by the addition of a water molecule (Fig. 2) (Zechel and Withers 2000). By contrast, lyases are assigned to polysaccharide lyase (PL) families based on sequence similarities (Lombard et al. 2010), which cleave the linkage between hexosamine and uronic acid through a elimination reaction to yield an unsaturated 4,5-bond between C-4 and C-5 on the uronic acid residue (Fig. 2) (Garron and Cygler 2010). GAG hydrolases and lyases are usually found in animals and microorganisms (Table 1), respectively.

2.1 Hyaluronidases (HAases)

HAases are a class of enzymes that can degrade HA and, in most cases, can digest low sulfated chondroitin sulfate and chondroitin with slower activity (Stern and Jedrzejas 2006). The HAases are usually divided into three types (Kreil 1995). The first group are hyaluronate-4-glycanohydrolases (EC 3.2.1.35) that function as endo-$β$-N-acetylhexosaminidase and digest HA mainly to a tetrasaccharide (GlcUA-GlcNAc-GlcUA-GlcNAc) (Fig. 3). This type of enzyme is always found in mammalian spermatozoa, lysosomes and insects or snake venoms. The second group are hyaluronate-3-glycanohydrolases (EC 3.2.1.36) that act as endo-$β$-D-glucuronidase and degrade HA mainly to a tetrasaccharide (GlcNAc-GlcUA-GlcNAc-GlcUA) (Fig. 3).
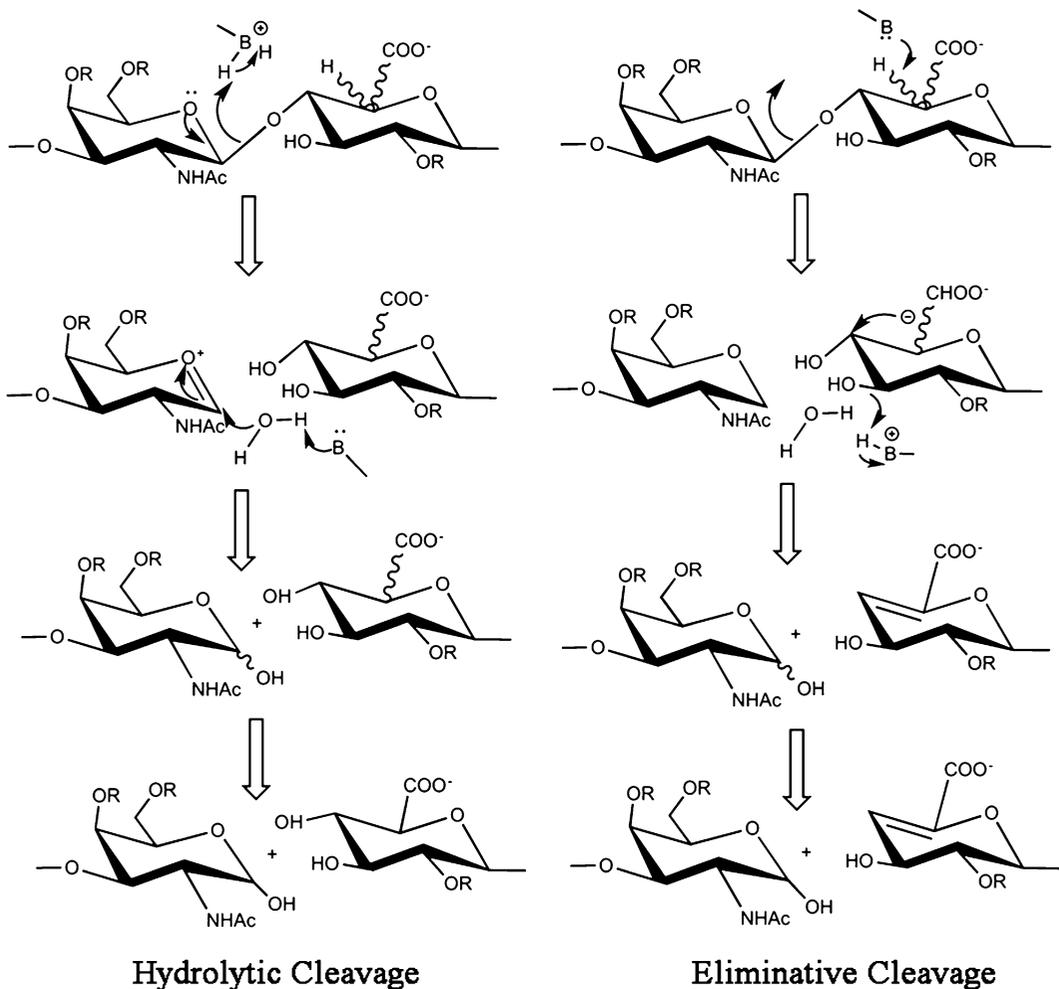
Fig. 2 Mechanism for GAG degraded by enzymes

This type of enzymes exist in some leeches and hookworms (Yuki and Fishman 1963). The last group is bacterial HA lyase (EC 4.2.2.1) and acts as N-acetylhexosaminidases via a β-elimination reaction across the L-1-4 linkage with introduction of an unsaturated bond 4,5-bond between C-4 and C-5 on the uronic acid residue (Garron and Cygler 2010). Unlike other groups of HAases, the final product of HA lyase is a disaccharide rather than a tetrasaccharide (Fig. 3).

2.1.1 Endo-β-N-Acetylhexosaminidase

All vertebrate HAases (EC 3.2.1.35) play a hydrolysis role in HA/CS digestion, and they also have transglycosidase activities (Hoffman

et al. 1956). Six known genes coding for HAase-like sequences have been found in the human genome, and all of them show a high degree of homology (Stern and Jedrzejas 2006; Csoka et al. 2001). They include human HYAL-1, −2, −3, −4 and PH20 as well as a pseudo gene HYAL-Phyal1 that is transcribed in humans but is not translated (Stern and Jedrzejas 2006). The first three human HAase genes (HYAL-1, −2 and −3) are clustered in the chromosome 3p21.3 locus, whereas the latter three genes are similarly clustered on chromosome 7q31.3 (Csoka et al. 2001; Stern 2003). Based on their optimal pH, HAases are divided into two categories. Most of them are considered as acidic

**Table 1** Properties and reaction conditions of enzymes acting on GAGs

Name	Substrate	Source	Degradation mechanism	Action pattern
HYAL-1 (Franzmann et al. 2003)	HA and CS	Human	Hydrolysis	Endo
HYAL-2 (Monzon et al. 2010)	HA	Human	Hydrolysis	Endo
HYAL-3 (Reese et al. 2010)	HA	Human	Hydrolysis	Endo
HYAL-4 (Kaneiwa et al. 2010)	CS	Human	Hydrolysis	Endo
PH20 (Cherr et al. 2001)	HA and CS	Human	Hydrolysis	Endo
CSase ABC I (Hamai et al. 1997)	HA, CS and DS	*Proteus vulgaris*	Lysis	Endo
CSase ABC II (Hamai et al. 1997)	HA, CS and DS	*Proteus vulgaris*	Lysis	Exo (from non-reducing end)
CSase AC I (Gu et al. 1995)	HA and CS	*Flavobacterium heparinum*	Lysis	Endo
CSase AC II (Yin et al. 2016; Lunin et al. 2004)	HA and CS	*Arthrobacter aurescens*	Lysis	Exo (from reducing end)
CSase B (Gu et al. 1995)	DS	*Flavobacterium heparinum*	Lysis	Endo
HAase-B (Guo et al. 2014)	HA and CS	*Bacillus* sp. A50	Lysis	Endo
P. aphidis HAase (Smirnou et al. 2015)	HA	*Pseudozyma aphidis*	Hydrolysis	Endo
C. laurentii HAase (Smirnou et al. 2015)	HA	*Cryptococcus laurentii*	Lysis	Endo
AcODV-E66 (Sugiura et al. 2011)	HA and CS	*Autographa californica* nucleopolyhedrovirus	Lysis	Endo
BmODV-E66 (Sugiura et al. 2013)	HA and CS	*Bombyx mori* nucleopolyhedrovirus	Lysis	Endo
HCLase (Han et al. 2014)	HA and CS	*Vibrio* sp. FC509	Lysis	Endo
HCDLase (Unpublished data)	HA, CS and DS	*Vibrio* sp. FC509	Lysis	Exo (from reducing end)
BniHL (Kurata et al. 2015)	HA and CS	*Bacillus niacin*	Lysis	Endo
ChoA1 (Kale et al. 2015)	HA and CS	*Arthrobacter* sp. MAT3885	Lysis	Endo

HAases because they have the highest activity at acidic pH (Lokeshwar et al. 2001). By contrast, PH20 is a neutral active HAase, as it is active at a neutral pH (Franzmann et al. 2003).

Among the six mammalian HAases, HYAL-1, −2 and PH20 are well characterized, and human HYAL-1 and HYAL-2 are two major HAases. HYAL-1 is a serum HAase, but its concentration in human serum is low (60 ng/ml) (Stern and Jedrzejas 2006). HYAL-1 can degrade high molecular weight HA to small oligomers with a random cut pattern and then subsequently to tetrasaccharides. Different from HYAL-1, HYAL-2 digests HA into ~20 kDa oligosaccharide fragments. HYAL-2 and PH20 are glycosyl phosphatidyl-inositol (GPI)-linked proteins that are necessary for ovum fertilization, which is why HAase can be used as a contraceptive (Garg et al. 2005a, b; Hardy et al. 2004; Suri 2004). Similar to most other HAases, the so-called HAases in humans also show a certain degree of CS-degrading activity, and some of them are considered to be CS- rather than HA-degrading enzymes. PH20 acts primarily on HA, but it can also degrade CS with lower activity. By contrast, HYAL-1 has recently been shown to hydrolyze CS-A more rapidly than HA (Yamada 2015). Furthermore, HYAL-4 is a misnomer as it has complete specificity for CS, with no degrading ability for HA (Csoka et al. 2001; Stern 2003; Jedrzejas and Stern 2005).

Fig. 3 Digestion of HA by different type of HAases
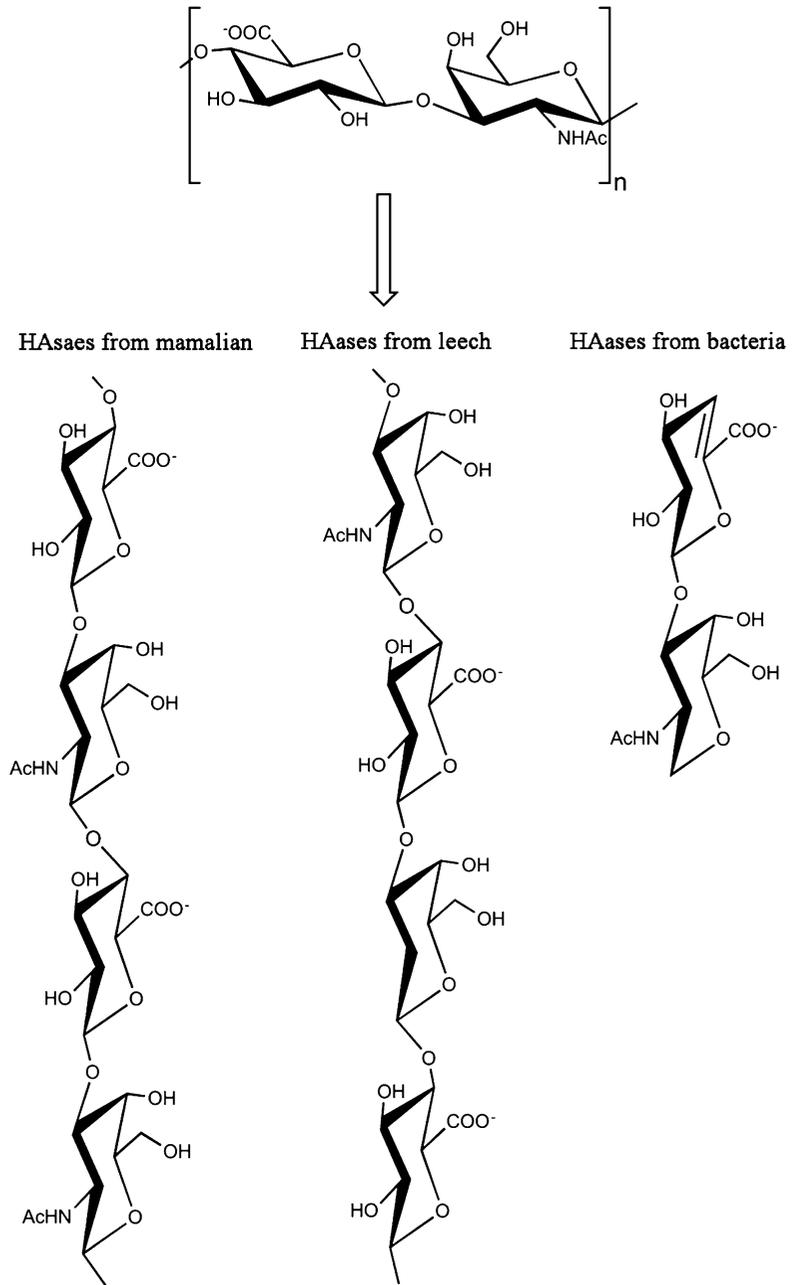
2.1.2 Endo-β-D-Glucuronidase

This type of enzyme, named annelids HAase, is the endo-β-D-glucuronidase (EC 3.2.1.36) that cleaves the β1-3 glycosidic bond between GlcUA and GlcNAc residues in HA chains. This type of enzyme has yet to be studied thoroughly and, thus, is only briefly mentioned here. Annelids HAase, usually found in leeches and some crustaceans (Hovingh and Linker 1999; Linker et al. 1957, 1960; Karlstam et al. 1991), degrades HA via a hydrolysis mechanism, which is similar to vertebrate HAase. However, the difference in the degradation pattern between vertebrate HAase and annelids HAase remains to be revealed by the study of their catalytic mechanisms.

2.1.3 HA Lyase

Bacterial HAase (EC 4.2.2.1), usually named HA eliminase or HA lyase, depolymerizes HA by β-elimination reaction with the introduction of an unsaturated bond (Li and Jedrzejas 2001; Kelly et al. 2001). There are a wide variety of bacteria that can produce HAase, such as *Proteus*, *Bacteroides fragilis*, *Streptococcus*, *Staphylococcus*, *Peptostreptococcus*, *Propionibacterium*, *Streptomyces*, *Clostridium*, and *Vibrio* (Han et al. 2014; Hynes and Walton 2000). Some of these lyases also have chondroitin lyase activity, for instance, those from *Aeromonas*, *Vibrio*, *Beneckea, Peptostreptococcus* and *Proteus*. In addition to the bacteria mentioned above, *Bacteroides* and *Fusobacterium* are also reported to produce HAase (Linhardt et al. 1987). HAases derived from Gram-negative bacteria are periplasmic enzymes, and they are less likely to play a role in pathogenesis. By contrast, most Gram-positive bacteria HAases are able to cause infections through skin abrasions of animals or humans (Hynes and Walton 2000). Many bacterial HAases digest HA via initial non-progressive endolytic activity, followed by exolytic degradation with the generation of unsaturated disaccharides as final products (Hovingh and Linker 1999; Jedrzejas et al. 2002). However, the digestion of CS by these so-called HAases is only via an endolytic action with unsaturated CS-disaccharides as final products.

2.2 Chondroitinases (CSases)

In animals, no enzyme family specifically degrading CS/DS has been identified so far. Based on recent studies, animal CS-degrading enzymes are categorized into the HAase family, which has been introduced above; thus, this section will mainly discuss CS/DS lyases from bacteria. Depending on their degradation pattern, CSases can be classified into two types: endo- and exo-lyases. The former digests GAG chains initially into larger oligosaccharides and finally to disaccharides with a random cut pattern, whereas the latter successively splits off disaccharide residues from the end of the sugar chain and does not produce any larger oligosaccharides in the process. Indeed, as in the case of other polysaccharide-degrading enzymes, usually both endo- and exo-enzymes for the same GAG substrate exist in the same genome, which should be a result of evolution for more effective degradation and exploitation of GAGs by bacteria. On the other hand, CSases can also be subdivided into three types based on their substrate, as discussed below.

2.2.1 CSase ABC

Lyases classified as CSase ABC can degrade both CS/DS and HA, irrespective of their sulfation/5-epimerization pattern. CSase ABC has been found in *Proteus vulgaris* (Yamagata et al. 1968) and *Bacteroides thetaiotaomicron* (Linn et al. 1983). These lyases from bacteria belong to the PL family 8 (www.cazy.org). Now, the most commonly used CSase ABC for GAG structure analysis and for the potential application as a therapeutic tool is CSase ABC from *Proteus vulgaris* (Fig. 4). As is known, CSase ABC from *Proteus vulgaris* is not one enzyme, but is a mixture of CSase ABC I (4.2.2.20) with endolytic activity and CSase ABC II (4.2.2.21) with exolytic activity. When degrading CS/DS, CSase ABC I catalyzes the endolytic cleavage of GAG chains that produce tetrasaccharides and disaccharides as final products. In contrast, CSase ABC II degrades GAG chains from the non-reducing end toward the reducing end (Hamai et al. 1997). CSase ABC I cannot degrade the GAG completely and thus works together with CSase ABC II as needed. The commercially available CSase ABC is a mixture of CSase ABC I and CSase ABC II for this reason.

2.2.2 CSase AC

Although as in the case of CSase ABC, CSase AC also belongs to the PL family 8 (www.cazy.org), CSase AC (EC 4.2.2.5) is highly sensitive to the 5-epimerization of GlcUA residues in GAG chains and can only act on CS, HA and CS domains in CS-DS hybrid chains (Linhardt et al. 1987, 2006). There are two well-known,
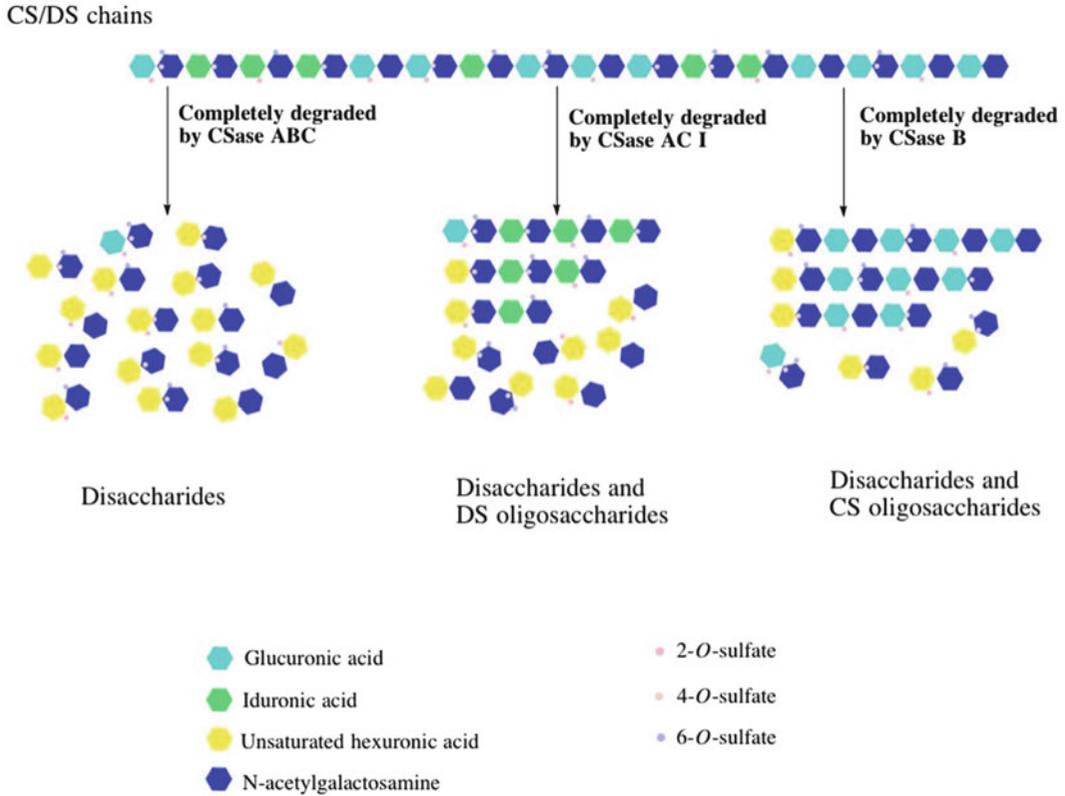
CS/DS chains

Fig. 4 The degradation profiles of CS/DS hybrid chain by different CSases

commercially available enzymes, CSase AC I from *Flavobacterium heparinum* and CSase AC II from *Arthrobacter aurescens*, that show endolytic and exolytic activities, respectively (Hiyama and Okada 1975). As an exo-type lyase, CSase AC II has been shown to digest CS and HA chains from the reducing end toward the non-reducing end (Zhang et al. 2008; Yin et al. 2016). In the structural and functional analyses of CS/DS chains, CSase AC I is a very useful tool for studying the disaccharide composition of CS domains as well as the distribution and preparation of DS domains in CS/DS heteromers (Fig. 4) (Li et al. 2007, 2010). By contrast, CSase AC II has great potential for the enzymatic sequencing of CS oligosaccharides due to its exolytic activity.

2.2.3 CSase B

CSase B (EC 4.2.2.19) from *Flavobacterium heparinum* is the only identified enzyme that is specific for the cleavage of DS and DS domains in CS-DS hybrid chains (Yamagata et al. 1968). Different from CSase ABC and CSase AC, CSase B belongs to the PL family 6 (www. cazy.org) with alginate lyase. CSase B has endolytic activity and can be used to investigate the disaccharide composition of DS domains and the distribution and isolation of CS domains in CS/DS chains (Fig. 4) (Li et al. 2007, 2010; Tkalec et al. 2000).

3   New Sources of HA/CS/DS Depolymerases

Most recently, some unique CS/DS/HA-degrading enzymes were identified from various microorganisms. Guo et al. found a novel HAase from *Bacillus* sp. A50, which was a lyase with extra high activity toward both HA and CS and was named HAase-B (Guo et al. 2014). Smirnou

et al. found two types of HAase from yeasts (Smirnou et al. 2015). The first one was *P. aphidis* HAase from *Pseudozyma aphidis*, which could hydrolyze the β-1,4 glycosidic bonds of HA to produce a range of even-numbered oligosaccharides, similar to vertebrate and venom HAases. The other one was *C. laurentii* HAase from *Cryptococcus laurentii*, which is a typical HA lyase and cleaves β-1,4 glycosidic bonds of HA via *β*-elimination reactions to produce unsaturated oligosaccharides. The laboratory of Watanabe identified two highly homologous occlusion-derived virus envelope protein 66s from *Autographa californica* nucleopolyhedrovirus and *Bombyx mori* nucleopolyhedrovirus, and both of them had specific lyase activity for non-sulfated and 6-*O*-sulfated CS (Sugiura et al. 2011, 2013).

Although marine animals are rich in GAGs with unique structures, no GAG degrading enzymes had been identified from marine organisms before HCLase was found in *Vibrio* sp. FC509. *Vibrio* sp. FC509 with ultra-highly degrading activity toward CS/DS was found and isolated by our laboratory from coastal sediments. Most GAG-degrading enzymes in the genome of this bacterium show a series of unique features, such as HCLase and HCDLase. HCLase has very high endolytic activity for HA and CSs with various sulfation patterns at an approximately neutral pH and exhibits excellent biochemical characteristics, such as pH and thermal stability and halophilicity (Han et al. 2014). HCDLase is a novel exo-type lyase that can degrade HA, CS and DS from their reducing end to the non-reducing end. In particular, this enzyme can effectively cut various sulfated oligosaccharides labeled with fluorescent dye, which makes it a very useful tool for oligosaccharide sequencing (unpublished data). Following our study, several other enzymes with HA/CS-degrading activity were identified from sea microorganisms by other groups. Kurata et al. found BniHL from a deep-sea bacteria *Bacillus niacin,* which has HA and CS activity at an approximately neutral pH (Kurata et al. 2015). Kale et al. found an enzyme similar to CSase AC I, named ChoA1, from *Arthrobacter*

sp. MAT3885, that can degrade chondroitin, CS and HA (Kale et al. 2015). Taken together, these studies indicate that the ocean is a huge untapped treasure trove for novel GAG-degrading enzymes.

4   Confusion Between "HAase" and "CSase"

As discussed above, an increasing number of studies have shown that the terms "HAase" and "CSase" are somewhat of misnomers because most animal and bacterial enzymes possess both HA and CS degradation activities. Some researchers think the reason for this overlap is because chondroitin is a precursor of HA in evolution (Stern and Jedrzejas 2006). For example, *Caenorhabditis elegans* contains only chondroitin (not CS) and one HA-like sequence but not HA (Yamada et al. 1999). The difference in structure between chondroitin and HA is that the GalNAc of chondroitin is replaced by the GlcNAc of HA. GalNAc and GlcNAc are isomers with epimerization at the C-4 site. In addition, HAase cannot digest Hep/HS, although the repeating disaccharides of HA and Hep/HS are composed of the same hexuronic acid (GlcUA) and hexosamine (GlcNAc/GlcN). Therefore, the composition of the disaccharide unit is not the key reason for the substrate specificity of enzymes. In comparison, the more crucial reason that causes structural differences of GAGs is the type of glycosidic bonds between monosaccharide residues. As is known, HA and CS have the same bond types (-4HexUAβ1-3HexNβ1-) between hexuronic acid (HexUA) and hexosamine (HexN) residues, which, however, are very different from those (-4HexUAβ/α1-4HexNα1-) of Hep/HS. This reason may be why most "HAases" and "CSases" can recognize and digest both HA and CS but not Hep/HS. In fact, to date, no enzyme has been found to possess both HAase/CSase and heparinase activity. To avoid a misnomer, therefore, the GAG-degrading enzymes can be named by the type of glycosidic bonds they recognize, or be simply named according to their substrate

specificity and enzymatic mechanism such as HCLase and HCDLase mentioned above.

5   Conclusions

Structural heterogeneity has hampered detailed structural and functional analyses of GAG chains. HA and CS/DS depolymerases with specific activities are very important tools for HA- and CS/DS-related research and applications. However, there are only a few HAases and CSases that have been characterized in detail and are commercially available so far, which should be due to the lack of powerful research means and new enzyme resources in the past. Therefore, it is now highly necessary to systematically characterize old enzymes with unclear activity and to search for novel enzymes from new resources. By reviewing the present status of studies of both HAases and CSases, we hope that this review can provide an overview of HA and CS/DS-degrading enzymes to the readers, and initiate further research in this field.

References

Bergefall K, Trybala E, Johansson M, Uyama T, Naito S, Yamada S, Kitagawa H, Sugahara K, Bergström T (2005) Chondroitin sulfate characterized by the E-disaccharide unit is a potent inhibitor of herpes simplex virus infectivity and provides the virus binding sites on gro2C cells. J Biol Chem 280(37):32193–32199

Boneu B (1995) Glycosaminoglycans: clinical use. In: Seminars in thrombosis and hemostasis, 1995. vol 2. p 209–212

Bülow HE, Hobert O (2006) The molecular diversity of glycosaminoglycans shapes animal development. Annu Rev Cell Dev Biol 22:375–407

Capila I, Linhardt RJ (2002) Heparin–protein interactions. Angew Chem Int Ed Engl 41(3):390–412

Cheng F, Heinegård D, Malmström A, Schmidtchen A, Yoshida K, Fransson L-A (1994) Patterns of uronosyl epimerization and 4-/6–0-sulphation in chondroitin/dermatan sulphate from decorin and biglycan of various bovine tissues. Glycobiology 4(5):685–696

Cherr GN, Yudin AI, Overstreet JW (2001) The dual functions of GPI-anchored PH-20: hyaluronidase and intracellular signaling. Matrix Biol 20(8):515–525

Clement AM, Nadanaka S, Masayama K, Mandl C, Sugahara K, Faissner A (1998) The DSD-1 carbohydrate epitope depends on sulfation, correlates with chondroitin sulfate D motifs, and is sufficient to promote neurite outgrowth. J Biol Chem 273(43):28444–28453

Csoka AB, Frost GI, Stern R (2001) The six hyaluronidase-like genes in the human and mouse genomes. Matrix Biol 20(8):499–508

Esko JD, Selleck SB (2002) Order out of chaos: assembly of ligand binding sites in heparan sulfate. Annu Rev Biochem 71(1):435–471

Faissner A, Clement A, Lochter A, Streit A, Mandl C, Schachner M (1994) Isolation of a neural chondroitin sulfate proteoglycan with neurite outgrowth promoting properties. J Cell Biol 126(3):783–799

Franzmann EJ, Schroeder GL, Goodwin WJ, Weed DT, Fisher P, Lokeshwar VB (2003) Expression of tumor markers hyaluronic acid and hyaluronidase (HYAL1) in head and neck tumors. Int J Cancer 106(3):438–445

Fraser J, Laurent T, Laurent U (1997) Hyaluronan: its nature, distribution, functions and turnover. J Intern Med 242(1):27–33

Garg A, Anderson R, Zaneveld L, Garg S (2005a) Biological activity assessment of a novel contraceptive antimicrobial agent. J Androl 26(3):414–421

Garg S, Vermani K, Garg A, Anderson RA, Rencher WB, Zaneveld LJ (2005b) Development and characterization of bioadhesive vaginal films of sodium polystyrene sulfonate (PSS), a novel contraceptive antimicrobial agent. Pharm Res 22(4):584–595

Garron M-L, Cygler M (2010) Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. Glycobiology 20(12):1547–1573

Gu K, Linhardt R, Laliberte M, Zimmermann J (1995) Purification, characterization and specificity of chondroitin lyases and glycuronidase from *Flavobacterium heparinum*. Biochem J 312(2):569–577

Guo X, Shi Y, Sheng J, Wang F (2014) A novel hyaluronidase produced by *Bacillus* sp. A50. PLoS One 9(4): e94156

Hamai A, Hashimoto N, Mochizuki H, Kato F, Makiguchi Y, Horie K, Suzuki S (1997) Two distinct chondroitin sulfate ABC lyases an endoeliminase yielding tetrasaccharides and an exoeliminase preferentially acting on oligosaccharides. J Biol Chem 272(14):9123–9130

Han W, Wang W, Zhao M, Sugahara K, Li F (2014) A novel eliminase from a marine bacterium that degrades hyaluronan and chondroitin sulfate. J Biol Chem 289(40):27886–27898

Handel T, Johnson Z, Crown S, Lau E, Sweeney M, Proudfoot A (2005) Regulation of protein function

by glycosaminoglycans-as exemplified by chemokines. Annu Rev Biochem 74:385–410

Hardy CM, Clydesdale G, Mobbs KJ, Pekin J, Lloyd ML, Sweet C, Shellam GR, Lawson MA (2004) Assessment of contraceptive vaccines based on recombinant mouse sperm protein PH20. Reproduction 127(3):325–334

Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 280(2):309–316

Hiyama K, Okada S (1975) Crystallization and some properties of chondroitinase from *Arthrobacter aurescens*. J Biol Chem 250(5):1824–1828

Hoffman P, Meyer K, Linker A (1956) Transglycosylation during the mixed digestion of hyaluronic acid and chondroitin sulfate by testicular hyaluronidase. J Biol Chem 219(2):653–663

Hovingh P, Linker A (1999) Hyaluronidase activity in leeches (Hirudinea). Comp Biochem Physiol B Biochem Mol Biol 124(3):319–326

Hsiao J-C, Chung C-S, Chang W (1999) Vaccinia virus envelope D8L protein binds to cell surface chondroitin sulfate and mediates the adsorption of intracellular mature virions to cells. J Virol 73(10):8750–8761

Hwang H-Y, Olson SK, Esko JD, Horvitz HR (2003) Caenorhabditis elegans early embryogenesis and vulval morphogenesis require chondroitin biosynthesis. Nature 423(6938):439–443

Hynes WL, Walton SL (2000) Hyaluronidases of Gram-positive bacteria. FEMS Microbiol Lett 183(2):201–207

Izumikawa T, Kitagawa H, Mizuguchi S, Nomura KH, Nomura K, J-i T, Gengyo-Ando K, Mitani S, Sugahara K (2004) Nematode chondroitin polymerizing factor showing cell-/organ-specific expression is indispensable for chondroitin synthesis and embryonic cell division. J Biol Chem 279(51):53755–53761

Jackson RL, Busch SJ, Cardin AD (1991) Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. Physiol Rev 71(2):481–539

Jedrzejas MJ, Stern R (2005) Structures of vertebrate hyaluronidases and their unique enzymatic mechanism of hydrolysis. Proteins 61(2):227–238

Jedrzejas MJ, Mello LV, de Groot BL, Li S (2002) Mechanism of hyaluronan degradation by *Streptococcus pneumoniae* hyaluronate lyase STRUCTURES OF COMPLEXES WITH THE SUBSTRATE. J Biol Chem 277(31):28287–28297

Kale V, Friðjónsson Ó, Jónsson JÓ, Kristinsson HG, Ómarsdóttir S, Hreggviðsson GÓ (2015) Chondroitin lyase from a marine *Arthrobacter* sp. MAT3885 for the production of chondroitin sulfate disaccharides. Mar Biotechnol (NY) 17(4):479–492

Kamhi E, Joo EJ, Dordick JS, Linhardt RJ (2013) Glycosaminoglycans in infectious disease. Biol Rev Camb Philo Soc 88(4):928–943

Kaneiwa T, Mizumoto S, Sugahara K, Yamada S (2010) Identification of human hyaluronidase-4 as a novel

chondroitin sulfate hydrolase that preferentially cleaves the galactosaminidic linkage in the trisulfated tetrasaccharide sequence. Glycobiology 20(3):300–309

Karlstam B, Vincent J, Johansson B, Brynö C (1991) A simple purification method of squeezed krill for obtaining high levels of hydrolytic enzymes. Prep Biochem 21(4):237–256

Kelly SJ, Taylor KB, Li S, Jedrzejas MJ (2001) Kinetic properties of *Streptococcus pneumoniae* hyaluronate lyase. Glycobiology 11(4):297–304

Klüppel M, Wight TN, Chan C, Hinek A, Wrana JL (2005) Maintenance of chondroitin sulfation balance by chondroitin-4-sulfotransferase 1 is required for chondrocyte development and growth factor signaling during cartilage morphogenesis. Development 132(17):3989–4003

Kreil G (1995) Hyaluronidases—a group of neglected enzymes. Protein Sci 4(9):1666–1669

Kurata A, Matsumoto M, Kobayashi T, Deguchi S, Kishimoto N (2015) Hyaluronate lyase of a deep-sea *Bacillus niacini*. Mar Biotechnol (NY) 17(3):277–284

Kusche-Gullberg M, Kjellén L (2003) Sulfotransferases in glycosaminoglycan biosynthesis. Curr Opin Struct Biol 13(5):605–611

Li S, Jedrzejas MJ (2001) Hyaluronan binding and degradation by *Streptococcus agalactiae* hyaluronate lyase. J Biol Chem 276(44):41407–41416

Li F, Shetty AK, Sugahara K (2007) Neuritogenic activity of chondroitin/dermatan sulfate hybrid chains of embryonic pig brain and their mimicry from shark liver INVOLVEMENT OF THE PLEIOTROPHIN AND HEPATOCYTE GROWTH FACTOR SIGNALING PATHWAYS. J Biol Chem 282(5):2956–2966

Li F, Nandini CD, Hattori T, Bao X, Murayama D, Nakamura T, Fukushima N, Sugahara K (2010) Structure of pleiotrophin-and hepatocyte growth factor-binding sulfated hexasaccharide determined by biochemical and computational approaches. J Biol Chem 285(36):27673–27685

Linhardt RJ, Toida T (2004) Role of glycosaminoglycans in cellular communication. Acc Chem Res 37(7):431–438

Linhardt R, Galliher P, Cooney C (1987) Polysaccharide lyases. Appl Biochem Biotechnol 12(2):135–176

Linhardt RJ, Avci FY, Toida T, Kim YS, Cygler M (2006) CS lyases: structure, activity, and applications in analysis and the treatment of diseases. Adv Pharmacol 53:187–215

Linker A, Hoffman P, Meyer K (1957) The hyaluronidase of the leech: an endoglucuronidase. Nature 180(4590):810–811

Linker A, Meyer K, Hoffman P (1960) The production of hyaluronate oligosaccharides by leech hyaluronidase and alkali. J Biol Chem 235(4):924–927

Linn S, Chan T, Lipeski L, Salyers AA (1983) Isolation and characterization of two chondroitin lyases from

Bacteroides thetaiotaomicron. J Bacteriol 156(2):859–866

Lokeshwar VB, Rubinowicz D, Schroeder GL, Forgacs E, Minna JD, Block NL, Nadji M, Lokeshwar BL (2001) Stromal and epithelial expression of tumor markers hyaluronic acid and HYAL1 hyaluronidase in prostate cancer. J Biol Chem 276(15):11922–11932

Lombard V, Bernard T, Rancurel C, Brumer H, Coutinho PM, Henrissat B (2010) A hierarchical classification of polysaccharide lyases for glycogenomics. Biochem J 432(3):437–444

Lunin VV, Li Y, Linhardt RJ, Miyazono H, Kyogashima M, Kaneko T, Bell AW, Cygler M (2004) High-resolution crystal structure of *Arthrobacter aurescens* chondroitin AC lyase: an enzyme–substrate complex defines the catalytic mechanism. J Mol Biol 337(2):367–386

Maccarana M, Olander B, Malmström J, Tiedemann K, Aebersold R, Lindahl U, J-p L, Malmström A (2006) Biosynthesis of dermatan sulfate CHONDROITIN-GLUCURONATE C5-EPIMERASE IS IDENTICAL TO SART2. J Biol Chem 281(17):11560–11568

Meyer K, Palmer JW (1934) The polysaccharide of the vitreous humor. J Biol Chem 107(3):629–634

Mizuguchi S, Uyama T, Kitagawa H, Nomura KH, Dejima K, Gengyo-Ando K, Mitani S, Sugahara K, Nomura K (2003) Chondroitin proteoglycans are involved in cell division of Caenorhabditis elegans. Nature 423(6938):443–448

Mizumoto S, Sugahara K (2013) Glycosaminoglycans are functional ligands for receptor for advanced glycation end-products in tumors. FEBS J 280(10):2462–2470

Monzon ME, Fregien N, Schmid N, Falcon NS, Campos M, Casalino-Matsuda SM, Forteza RM (2010) Reactive oxygen species and hyaluronidase 2 regulate airway epithelial hyaluronan fragmentation. J Biol Chem 285(34):26126–26134

Nadanaka S, Clement A, Masayama K, Faissner A, Sugahara K (1998) Characteristic hexasaccharide sequences in octasaccharides derived from shark cartilage chondroitin sulfate D with a neurite outgrowth promoting activity. J Biol Chem 273(6):3296–3307

Nandi S, Akhter MP, Seifert MF, Dai X-M, Stanley ER (2006) Developmental and functional significance of the CSF-1 proteoglycan chondroitin sulfate chain. Blood 107(2):786–795

Reese KL, Aravindan RG, Griffiths GS, Shao M, Wang Y, Galileo DS, Atmuri V, Triggs-Raine BL, Martin-DeLeon PA (2010) Acidic hyaluronidase activity is present in mouse sperm and is reduced in the absence of SPAM1: evidence for a role for hyaluronidase 3 in mouse and human sperm. Mol Reprod Dev 77(9):759–772

Silbert JE, Sugumaran G (2002) Biosynthesis of chondroitin/dermatan sulfate. IUBMB life 54(4):177–186

Smirnou D, Krčmář M, Kulhánek J, Hermannová M, Bobková L, Franke L, Pepeliaev S, Velebný V (2015) Characterization of hyaluronan-degrading enzymes from yeasts. Appl Biochem Biotechnol 177(3):700–712

Stern R (2003) Devising a pathway for hyaluronan catabolism: are we there yet? Glycobiology 13(12):105R–115R

Stern R, Jedrzejas MJ (2006) Hyaluronidases: their genomics, structures, and mechanisms of action. Chem Rev 106(3):818–839

Sugahara K, Mikami T (2007) Chondroitin/dermatan sulfate in the central nervous system. Curr Opin Struct Biol 17(5):536–545

Sugahara K, Mikami T, Uyama T, Mizuguchi S, Nomura K, Kitagawa H (2003) Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. Curr Opin Struct Biol 13(5):612–620

Sugiura N, Setoyama Y, Chiba M, Kimata K, Watanabe H (2011) Baculovirus envelope protein ODV-E66 is a novel chondroitinase with distinct substrate specificity. J Biol Chem 286(33):29026–29034

Sugiura N, Ikeda M, Shioiri T, Yoshimura M, Kobayashi M, Watanabe H (2013) Chondroitinase from baculovirus Bombyx mori nucleopolyhedrovirus and chondroitin sulfate from silkworm Bombyx mori. Glycobiology 23(12):1520–1530. cwt082

Suri A (2004) Sperm specific proteins-potential candidate molecules for fertility control. Reprod Biol Endocrinol 2(10)

Taylor KR, Gallo RL (2006) Glycosaminoglycans and their proteoglycans: host-associated molecular patterns for initiation and modulation of inflammation. FASEB J 20(1):9–22

Tkalec AL, Fink D, Blain F, Zhang-Sun G, Laliberte M, Bennett DC, Gu K, Zimmermann JJ, Su H (2000) Isolation and expression in Escherichia coli ofcslA and cslB, genes coding for the chondroitin sulfate-degrading enzymes Chondroitinase AC and Chondroitinase B, respectively, from *Flavobacterium heparinum*. Appl Environ Microbiol 66(1):29–35

Williams RK, Straus SE (1997) Specificity and affinity of binding of herpes simplex virus type 2 glycoprotein B to glycosaminoglycans. J Virol 71(2):1375–1380

Yamada S (2015) Role of hyaluronidases in the catabolism of chondroitin sulfate. In: Biochemical roles of eukaryotic cell surface macromolecules. Springer, Heidelberg, pp 185–197

Yamada S, Van Die I, Van den Eijnden DH, Yokota A, Kitagawa H, Sugahara K (1999) Demonstration of glycosaminoglycans in Caenorhabditis elegans. FEBS Lett 459(3):327–331

Yamagata T, Saito H, Habuchi O, Suzuki S (1968) Purification and properties of bacterial chondroitinases and chondrosulfatases. J Biol Chem 243(7):1523–1535

Yin F-X, Wang F-S, Sheng J-Z (2016) Uncovering the catalytic direction of chondroitin AC exolyase: from the reducing end towards the non-reducing end. J Biol Chem 291(9):4399–4406

Yuki H, Fishman WH (1963) Purification and characterization of leech hyaluronic acid-endo-β-glucuronidase. J Biol Chem 238(5):1877–1879

Zechel DL, Withers SG (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst. Acc Chem Res 33(1):11–18

Zhang Z, Park Y, Kemp MM, Zhao W, Im A, Shaya D, Cygler M, Kim YS, Linhardt RJ (2008) Liquid chromatography–mass spectrometry to study chondroitin lyase action pattern. Anal Biochem 385:57–64

Factors that Control Mitotic Spindle Dynamics

Roberta Fraschini

Abstract

Mitosis is the last phase of the cell cycle and it leads to the formation of two daughter cells with the same genetic information. This process must occurr in a very precise way and this task is essential to preserve genetic stability and to maintain cell viability. Accurate chromosome segregation during mitosis is brought about by an important cellular organelle: the mitotic spindle. This structure is made of microtubules, polymers of alpha and beta tubulin, and it is highly dynamic during the cell cycle: it emanates from two microtubules organizing centers (Spindle Pole Bodies, SPBs, in yeast) that are essential to build a short bipolar spindle, and it undergoes two steps of elongation during anaphase A and anaphase B in order to separate sister chromatids. Several proteins are involved in the control of mitotic spindle dynamics and their activity is tightly coordinated with other cell cycle events and with cell cycle progression.

Abbreviations

APC	Anaphase Promoting Complex
CDK	cyclin dependent kinase
cMTs	cytoplasmic microtubules
CPCs	chromosome passenger complexes
EM	electron microscopy
FEAR	Cdc-Fourteen Early Anaphase Release pathway
kMTs	kinetochore MTs
MAP	microtubule-associated protein
MEN	Mitotic Exit Network
MTs	microtubules
nMTs	nuclear MTs
SAC	spindle assembly checkpoint
SPB	Spindle Pole Bodies

R. Fraschini (✉)
Dipartimento di Biotecnologie e Bioscienze, Università degli Studi di Milano-Bicocca, Piazza della Scienza 2, 20126 Milan, Italy
e-mail: roberta.fraschini@unimib.it

1 Introduction

The mitotic spindle is a highly dynamic structure and this characteristic is essential in order to perform its function. In yeast cells, like in animal cells, the mitotic spindle is formed by microtubules (MTs). One microtubule is a cilinder composed of 13 protofilaments that are made of heterodimers of α- and β-tubulin assembled together in a head-to-tail fashion. As a result, each microtubule has two distinct ends with different properties: a dynamic fast growing one (plus end) and a slow growing one (minus end). The passage from growth to shortening and vice versa is stochastic and this dynamic instability is important for MT attachment to the kinetochore and for chromosome segregation, in higher eukaryotes it has been observed that MTs grow and shrink continuously and their subunits move towards the pole, thus generating MT flux, important for spindle functionality. MTs polarity is very important for their function during mitosis. Indeed, during growing and dwindling, both pushing and pulling forces are generated and different microtubule-binding proteins show preference in binding plus or minus end, so microtubule polarity enables directional movement of motor proteins that are bound to MTs. Motor proteins move along the MTs and generate

forces that allow adiacent MTs to slide in relation to one another.

Microtubules are divided into different subsets: astral MTs (three for each SPB) that are responsible for interactions between the spindle and the cell cortex, kinetochore MTs (kMTs) (one for each centromere) that attach to the chromosomes and to the spindle poles and interpolar MTs (approximately four from each pole) that interdigitate between the two spindle poles to form an antiparallel microtubules array that is called "spindle midzone" (Fig. 1). MTs are responsible for spindle orientation and nuclear positioning during the cell cycle and for spindle dynamics during mitosis.

Every kind of MT is bound by microtubule-associated proteins (MAPs) and motors, proteins that convert chemical energy to mechanical energy (forces). It is now clear that motor proteins display plus-end and minus-end directed motility and also polymerase and depolymerase activity alone or in complex with interacting proteins. Budding yeast has several motors, described below. Cin8 and Kip1, members of kinesin-5 family, provide a pushing force that drives spindle elongation, this force is counteracted by Kar3, a minus-end directed motor member of kinesin-14 family (Roof et al. 1992; Saunders and Hoyt 1992). Kar3 is

Fig. 1 The α and β tubulin subunits assemble in mitotic spindle microtubules (MTs) that have a polarity: the minus (−) ends are located near the centrosomes (Spindle Pole Bodies, in yeast), while the plus (+) end is at the opposite direction and is where increase of MT lenght occurs
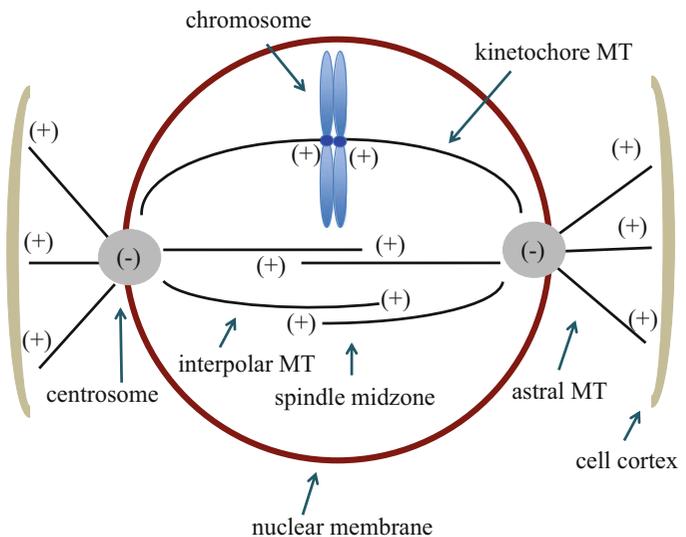
able to bind to microtubule plus ends and to function as a depolymerase (Molk and Bloom 2006), it is one of the regulators that promote transport of captured kinetochores along MTs. There is another protein that exhibit depolymerase activity: Kip3, a kinesin-8 family member that cooperatively disassociates from plus end microtubules, resulting in the removal of one or two tubulin dimers and that plays important roles in spindle dynamics (Gupta et al. 2006; Varga et al. 2006). In the nucleus, Kip3 helps chromosome congression in metaphase (Wargacki et al. 2010). In the cytoplasm, Kip3 contributes to position the spindle correctly at the bud neck in concert with some plus-end binding proteins such as Bim1, Bni1, Bud6 and Kar9 (Pearson and Bloom 2004). Kip2 is a cytoplasmatic kinesin that exhibit polymerization activity, it counteracts Kip3 action and helps spindle orientation with respect to mother bud axis, together with Bik1, the dynein Dyn1, Kar9 and Num1 (Carvalho et al. 2003, 2004). Dyn1 is a cytoplasmatic minus-end directed motor and plays an essential role in nuclear migration and mitotic spindle orientation, together with the dynactin complex, Bik1 and Kip2 (Lee et al. 2005; Moore et al. 2009). During spindle elongation it couples astral microtubule depolymerization to spindle pole movement.

Microtubule-associated proteins (MAPs) are proteins that bind MTs and participate in several processes, such as kMTs dynamics, spindle orientation and stability, regulation of motor protein function. Bim1, Bik1, Nip100, Stu1 and Stu2 bind MT plus ends. Bim1 plays an important role in spindle orientation (Hwang et al. 2003). Bik1 binds Kip2 and Bim1 and seems to promote Dyn1 function (Carvalho et al. 2004). Stu1, the yeast member of the CLASP family, binds β-tubulin and is required for SPB separation and spindle formation, it functions to stabilize MT plus ends by facilitating tubulin subunits incorporation, it is also present in the midregion of anaphase spindles at the plus ends of overlapping interpolar microtubules (Maiato et al. 2005; Pasqualone and Huffaker 1994; Yin et al. 2002). Stu2 is a dynamicity factor that promotes MTs polymerization at plus ends, it

localizes at microtubule plus ends and at kinetochores that are not yet bound to microtubules thus helping chromosome capture and it also contributes to mitotic spindle elongation during anaphase (Kosco et al. 2001; Severin et al. 2001). Nip100, the large subunit of dynactin complex, binds to dynein, it is involved in mitotic spindle positioning and mediates translocation of the mitotic spindle through the bud neck during anaphase (Kahana et al. 1998).

The spindle midzone is the central part of the spindle, where interpolar MTs interdigitate to form an antiparallel MTs array. The overlapping region is about 3–4 µm long and it recruits several proteins, such as kinesins Cin8/Kip1, Kar3, Bim1 and Ase1. The spindle midzone provides physical support from spindle formation to its disassembly and it is essential for spindle stability. The highly conserved Ase1 protein forms a dimer, binds MTs and crosslinks antiparallel MTs, thus contributing to spindle stability (Pellman et al. 1995), Ase1 is important for both spindle assembly and spindle elongation during anaphase (Schuyler et al. 2003; Kotwaliwale et al. 2007), see below. In anaphase, several proteins (CPCs, chromosome passenger complexes) are recruited to the spindle midzone in order to control spindle elongation and to inhibit abscission, the final step of cytokinesis, until chromosomes are completely segregated (Norden et al. 2006), see below.

In this review we will describe mitotic spindle dynamics during the cell cycle in the budding yeast *Saccharomyces cerevisiae*, a very useful model organism to dissect complex cellular processes. It is important to point out that all the studies on this topic have shown that several processes and key players are evolutionarily conserved from yeast to multicellular eukaryotes.

2 Chapter 1. Spindle Pole Bodies Separation and Bipolar Spindle Formation

The mitotic spindle is made of microtubules (MTs) that are nucleated from the the spindle pole bodies (SPBs), the functional equivalents

of the centrosomes. The SPB is inherited from the mother cell and duplicates once in each division cycle during early S phase; after the duplication, the two SPBs move away from each other and coordinate the MTs in order to assemble a short bipolar spindle before entry into mitosis. Budding yeast undergoes a "closed mitosis", i. e. the cells do not break down the nuclear envelope during mitosis, as a consequence the SPBs are embedded in the nuclear envelope throughout the yeast life cycle, they face both the nucleus and the cytoplasm and they are able to nucleate both nuclear (nMTs) and cytoplasmic microtubules (cMTs). cMTs play a key role in nuclear positioning since they drive nuclear movement towards the bud neck (Hildebrandt and Hoyt 2000) while nMTs form a structure that binds chromosomes and ensure their accurate segregation during mitosis.

The SPB is composed of at least 30 different proteins (Jaspersen and Winey 2004), 18 of which are core component, since they are always present in the structure importantly, 16 of these proteins are encoded by essential genes and 11 of them have vertebrate orthologs. SPB structure was determined using several structural analyses such as electron microscopy and electron tomography (Adams and Kilmartin 1999; Bullitt et al. 1997; O'Toole et al. 1999). It appears as a cylindrical, layered organelle that consists of three plaques: an outer plaque that faces the cytoplasm, an inner plaque that faces the nucleoplasm and a central plaque that is inside the nuclear membrane. The inner and outer plaques are the sites from which MTs are nucleated, consistently electron tomography showed that all microtubules ends at the SPBs are minus ends (O'Toole et al. 1999).

SPB duplication is an essential event for mitotic spindle formation, it is strictly coordinated with cell cycle progression and can be divided into three steps. Early in G1 a satellite, composed of core SPB components, is deposited next to the existing SPB. Then, after START (the point of commitment to the cell cycle), the satellite expands into a duplication plaque, a structure that is similar to the cytoplasmic half of a mature SPB. The last step is the insertion of the duplication plaque into the nuclear envelope through an opened nuclear pore and the addition of nuclear SPB components, such as Spc110p (Adams and Kilmartin 1999), this correlates with the acquisition of microtubule nucleation capacity. At the end of this process, cells contain two duplicated side-by-side SPBs that are different from each other since one is the older (mother) SPB and the other one is the younger SPB. Interestingly, the old SPB is already able to nucleate MTs while the new one is not (Segal et al. 2000; Pereira et al. 2001), this property allows the tethering of the old SPB with cortical landmarks in the bud and ensures that the old SPB migrates in the daughter cell. This asymmetry ensures the correct mitotic spindle orientation along the cell polarity axis, defined by the site of bud emergence. This task is essential in budding yeast cells that divide asymmetrically in order to ensure correct nuclear division between the mother and daughter cells. Establishment of spindle polarity occurs during spindle morphogenesis and relies on spindle pole body functionality and signals originating from the cell cortex that direct cytoplasmic microtubule attachments. Interestingly in several asymmetrically dividing cells the poles of the mitotic spindle do not segregate randomly between their two daughters. For example, during asymmetric cell divisions of neural progenitors in the mouse neocortex, the old centrosome is retained by progenitor cells and the new centrosome is acquired by differentiating cells (Wang et al. 2009). So the pattern of inheritance based on the intrinsic ability of the "old centrosome" to retain MT organization seems to be a general principle in asymmetric cell divisions.

The duplicated SPBs are connected by a bridge that must be severed to allow SPBs separation and proper mitotic spindle formation, this process also leads to the formation of two half bridge structures on which the satellite will form in the next cell cycle. Then the SPBs move apart thanks to microtubules, the activity of kinesin-like motor proteins Cin8 and Kip1 (Roof et al. 1992; Jacobs et al. 1988) and a high cyclin dependent kinase (CDK) activity (Haase et al. 2001).

The formation and maintenance of a bipolar spindle depends on a balance of forces acting on the spindle poles. Each SPB emanates 4 short nuclear MTs that have crossbridges that load motor protein needed to separate the SPBs, together with Ase1 and Stu1 (O'Toole et al. 1999). The primary outward force is generated by the plus end–directed kinesins Cin8 and Kip1 and it is counteracted by the minus-end-directed kinesin Kar3. Stu1 is a nonmotor protein that forms homodimers which cross-link antiparallel microtubules (Yin et al. 2002), Ase1 is a MAP that is phosphorylated by AuroraA/Ipl1 kinase and translocated to antiparallel MT to stabilize spindle midzone (Kotwaliwale et al. 2007). The proposed model suggests that Cin8 and Kip1 cross-link overlapping polar microtubules and slide them past one another to generate an outward force on the spindle. Ipl1/Ase1 pathway becomes essential for spindle assembly in the absence of Cin8 (Kotwaliwale et al. 2007).

Since the formation of a functional bipolar spindle is essential for cell viability, several pathways are involved in this process. Interestingly even cytoplasmic dynein and the actin cytoskeleton, which are involved in spindle positioning, are required for SPB separation in certain conditions, indicating that this process depends also on pulling forces on astral microtubules (Chiroli et al. 2009). Consonantly, cytoplasmic dynein is required for centrosome separation in mammalian cells (Vaisberg et al. 1993) and in *C. elegans* embryos (De Simone et al. 2016). In this recent publication the authors show that centrosome separation is powered by the joint action of dynein at the nuclear envelope and at the cell cortex.

Phenotypical analyses of mutants suggest that the force generated by bundling action is sufficient to break the intra-SPB bridge and the motor activity is not essential for this process (Crasta et al. 2006; Gheber et al. 1999). Cin8, Kip1 and Ase1 levels are regulated by proteolysis, in particular they are targeted for degradation by the E3 ubiquitin ligase Anaphase Promoting Complex (APC) associated to the regulatory subunit Cdh1. APC/Cdh1 activity is high in G1 phase of

the cell cycle and decreases as cell cycle proceeds, its inactivation allows Cin8, Kip1 and Ase1 accumulation, this in turn enables SPBs separation and assemble a short spindle during S phase (Crasta et al. 2006).

Cdh1 inactivation is very complex and involves the activity of the polo kinase Cdc5, the cyclin- dependent kinase Cdc28/Clb (Cdk1) and the Acm complex (Crasta et al. 2008). Cdh1 is partially inactivated through phosphorylation at multiple sites, in this process Cdc28 acts as a priming kinase: Cdh1 contains 11 Cdk1-phosphorylation sites and phosphorylation of these sites creates polo-box-binding (PBB) sites that allow the recruitment of the kinase Cdc5 to Cdh1 for further phosphorylation. In addition, physical binding of the Acm1/Bmh1/Bmh2 complex to Cdh1 forms an additional inhibitory pathway (Martinez et al. 2006). Thus, yeast cells use several pathways to inactivate the APC inhibitor Cdh1 which normally prevents accumulation of Cin8, Kip1 and Ase1, restraining cells to assemble a mitotic spindle. Complete inactivation of Cdh1 allows stabilization of microtubule binding proteins Cin8, Kip1 and Ase1 and mediates the formation of a short spindle (Crasta et al. 2008).

Centromeric DNA is replicated early during S phase and it is likely that kMTs are already able to bind kinetochores at this stage, thus forming a short pre-metaphase spindle. At the end of G2 phase, in yeast cells like in animal cells, a "metaphase" occurs and it is defined as a condition in which bipolar attachment is achieved before chromosome segregation. A metaphase spindle is 1–2 μm long and it is composed of two classes of MTs: core and kinetocore (Winey et al. 1995) (Fig. 2). The core or interpolar MTs keep the two SPBs in contact with each other and serve to push them apart during anaphase B. The kMTs are 0.4 μm long and connect the kinetochore with one SPB, likely a single microtubule is bound to the kinetochore, these MTs are essential to obtain a merotelic attachment that ensure a faithful chromosome segregation to daughter cells.

As soon as all of the chromosomes have bound microtubules emanating from either spindle pole through their kinetochores in order to achieve bi-orientation on the metaphase spindle,

Fig. 2 The metaphase
spindle is composed of core
MTs and kinetochore MTs.
Chromosomes are located
at the center of the spindle
thanks to the balance of
opposite forces: cohesin
that keeps sister chromatids
together and pulling forces
at the kinetochore MTs
(*red arrows*)
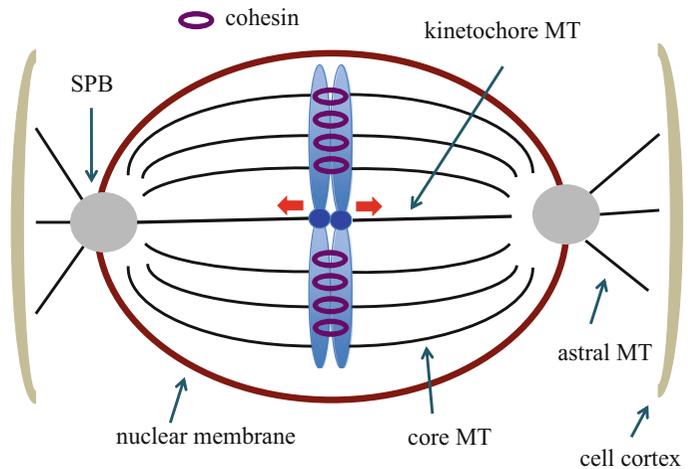
the spindle assembly checkpoint (SAC) is switched off, allowing the activation of the Anaphase Promoting Complex (APC) associated with its regulatory subunit Cdc20. This activates the protease separase which cleaves the cohesin complex that holds sister chromatids together, thereby starting anaphase (Clarke and Bachant 2008).

3   Chapter 2. Mitotic Spindle Dynamics During Anaphase A

At the metaphase to anaphase transition, the removal of proteins that hold sister chromatids together (cohesins) allows their separation. This is brought about by forces exerted by kinetochore-associated microtubules that pull them apart, indeed cohesin dissolution is not sufficient to ensure sister chromatids segregation. Anaphase onset causes a linear spindle elongation with about 1 μm/min speed.

The process of chromosome migration occurs in two steps: anaphase A, in which chromosomes move towards the spindle poles and anaphase B, in which the spindle poles move away from each other. Anaphase A is assumed to be dependent on shortening of the kinetochore microtubules that link centromeres to spindle poles. During anaphase A kinetochore MTs are shortened from 0.4 μm during metaphase to 0.2 μm in late anaphase spindles, this is probably achieved through

loss of tubulin dimers at each end of MTs (Fig. 3). Anaphase B involves pushing forces generated at the spindle midzone through motor proteins that cause the sliding of interdigitating microtubules emanating from the opposite spindle poles and pulling forces generated by motor proteins associated with astral microtubules that link the spindle to the cell cortex (Yeh et al. 1995).

The movement of chromosomes towards the pole during anaphase A is the result of the action of different mechanisms that likely occurr concomitantly. Studies in higher eukaryotes have revealed the existence of a "flux" that is the poleward movement of MTs, coupled to minus-end disassembly at the spindle pole (Maddox et al. 2003). During metaphase, plus-ends grow at kinetochores while minus-ends disassemble at poles, achieving a balance that maintains a constant spindle length. Inhibition of this flux leads to metaphase spindle colapse (Waters et al. 1996). During metaphase the flux exerts a constant poleward force on attached kinetochores and produces tension between sister kinetochores. At anaphase onset, after cohesin removal, this force and the cessation of kMT plus-end assembly lead to the segregation of sister chromatids to cell poles (Rogers et al. 2005). Likely kMTs are actively depolymerized from their kinetochore-associated plus ends and shortened at their spindle pole-associated minus ends, simultaneously. In
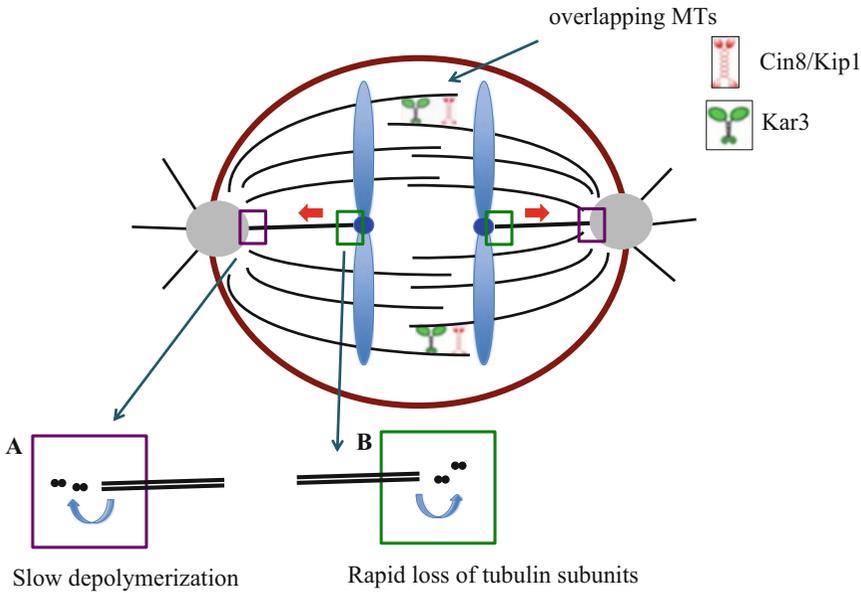
Fig. 3 During anaphase A chromosomes move towards the spindle poles thanks to pulling forces (*red arrows*) at the kinetochore microtubules (kMTs) that link centromeres to spindle poles. kMTs are shortened by slow depolymerization at their minus end (*A*) and by rapid loss of tubulin subunits at their plus end (*B*). In the midzone region, where interpolar MTs overlap, plus-end directed motors (Cin8 and Kip1) and minus-end directed kinesin Kar3 localize

addition, chromosome movement is governed also by plus-end-directed motors, such as members of kinesin 5 family, that produce a traction force by interacting with nearby kMTs to slide them poleward (Roof et al. 1992; Jacobs et al. 1988).

Mitotic spindle elongation requires a perfect balance between kinase an phosphatase activities indeed several proteins involved in spindle formation and elongation undergo phosphorylation events that are important for their function (Winey and Bloom 2012), at the metaphase to anaphase transition these modifications are reversed due to inactivation of mitotic CDKs and activation of phosphatases. Indeed, whereas metaphase is characterized by high Cdk1 activity, anaphase is marked by a reduction of Cdk1 activity, caused also by the activation of the phosphatase Cdc14. During most of the cell cycle Cdc14 is kept inactive inside the nucleolus and, at the anaphase onset, Cdc14 is released and activated in two steps. First, the Cdc-Fourteen Early Anaphase Release (FEAR) pathway triggers a partial and transient release of Cdc14 from the nucleolus to the nucleus (Stegmeier et al. 2002). FEAR-activated Cdc14 is involved in the control of microtubules dynamics by targeting several spindle associated proteins that form the spindle midzone and/or that are required for the stabilization of microtubules in anaphase cells (Higuchi and Uhlmann 2005; Woodbury and Morgan 2007). However, the FEAR-controlled Cdc14 is not sufficient to promote mitotic exit; only the full release of Cdc14 allows it, by counteracting and inhibiting Cdk1 activity. This full release is controlled by another regulatory pathway, the Mitotic Exit Network (MEN) that promotes mitotic exit and induces cytokinesis (Stegmeier and Amon 2004). These two functions are both controlled by Cdc14 and are largely independent, because cells can exit mitosis while failing to complete cytokinesis. In late anaphase, Cdc14 is able to dephosphorylate other proteins thus helping mitotic spindle elongation (Pereira and Schiebel 2003). The fact that cytokinesis initiation needs MEN activation ensures that cell division will not take place before anaphase is completed. The regulation of several

factors associated with the spindle by Cdc14 release provides a link between anaphase onset and the changes in microtubule dynamics that are required for the completion of mitotic spindle elongation.

As already mentioned, the spindle midzone is a crucial structure for spindle stabilization and elongation and it is also required to maintain the integrity of the anaphase spindle. Moreover, in *S. cerevisiae* spindle elongation is driven mostly by forces generated within the spindle midzone. Indeed, physical or genetic perturbations of the spindle midzone result in inefficient separation of the spindle poles and collapse of the anaphase spindle (Schuyler et al. 2003; Khodjakov et al. 2004).

FEAR-activated Cdc14 dephosphorylates several microtubule-associated proteins (MAPs) that reside at the spindle midzone and/or that are required for the stabilization of microtubules observed in anaphase cells. Among other factors, Cdc14 dephosphorylates the MAPs Ase1, Stu1 and Fin1, the chromosomal passenger protein Sli15, and the DASH kinetochore complex member Ask1; the dephosphorylations of all these proteins are required for proper midzone assembly and stability and/or for the microtubules dynamics in anaphase cells (Higuchi and Uhlmann 2005; Woodbury and Morgan 2007; Pereira and Schiebel 2003; Khmelinskii and Schiebel 2008; Khmelinskii et al. 2007, 2009; Rozelle et al. 2011). The regulation of midzone assembly by Cdc14 release connects anaphase onset with the assembly and function of a stable midzone.

Recently, we demonstrated that the kinase Swe1 plays an important and uncovered role in spindle dynamics and it is regulated by Cdc14 (Raspelli et al. 2015). It is well known that the protein kinase Swe1 inhibits entry into mitosis by phosphorylating the Y19 residue of the catalytic subunit of Cdk1, Cdc28 (Booher et al. 1993). Swe1 activity, levels and subcellular localization are finely regulated by phosphorylation and dephosphorylation events, ubiquitylation, degradation (Kellogg 2003; McMillan et al. 2002). At the end of G2 phase, Swe1 is degraded via the proteasome and this event allows mitotic entry

(McMillan et al. 2002). Little is known about Swe1 role after this point of the cell cycle although it is reported that a pool of Swe1 persists in the cells after mitotic entry (Raspelli et al. 2011; King et al. 2013), suggesting that it might play a role also in mitotic progression. However, to date, no other Swe1 substate has been identified.

We collected evidences that indicate that Swe1 could act as a mitotic spindle inhibitor independently of its action on Cdc28 (Raspelli et al. 2015). In addition we showed that Swe1 is phosphorylated and mostly localized in the nucleus and in the cytoplasm of the mother cell during metaphase then, during anaphase, it is dephosphorylated and spreaded also to the bud concomitantly with nuclear division. In agreement with published data (Breitkreutz et al. 2010), our further analysis showed that Cdc14 can bind Swe1 *in vivo* and that it is involved in its dephosphorylation. We hypothesize that Swe1 contributes to restrain mitotic spindle elongation until anaphase and that its inactivation by dephosphorylation is required to allow timely mitotic spindle elongation, thus ensuring the maintenance of genomic stability. An interesting question that remains unanswered is which protein is kept inactive or activated by Swe1 in order to block spindle elongation. Since we could never observe Swe1 localization at the spindle we hypothesize that Swe1 might form a complex with a diffusible protein that binds microtubules associated proteins or the spindle only after Swe1 inactivation.

4   Chapter 3. Mitotic Spindle Dynamics During Anaphase B

Anaphase B is the process that contributes most to the separation of sister chromatids, indeed a metaphase 1–2 μm spindle is elongated to 10 μm (Winey et al. 1995). During anaphase B the spindle poles move away from each other thanks to pushing forces generated at the spindle midzone through motor proteins that cause the sliding of interdigitating microtubules emanating from the opposite spindle poles. The full spindle
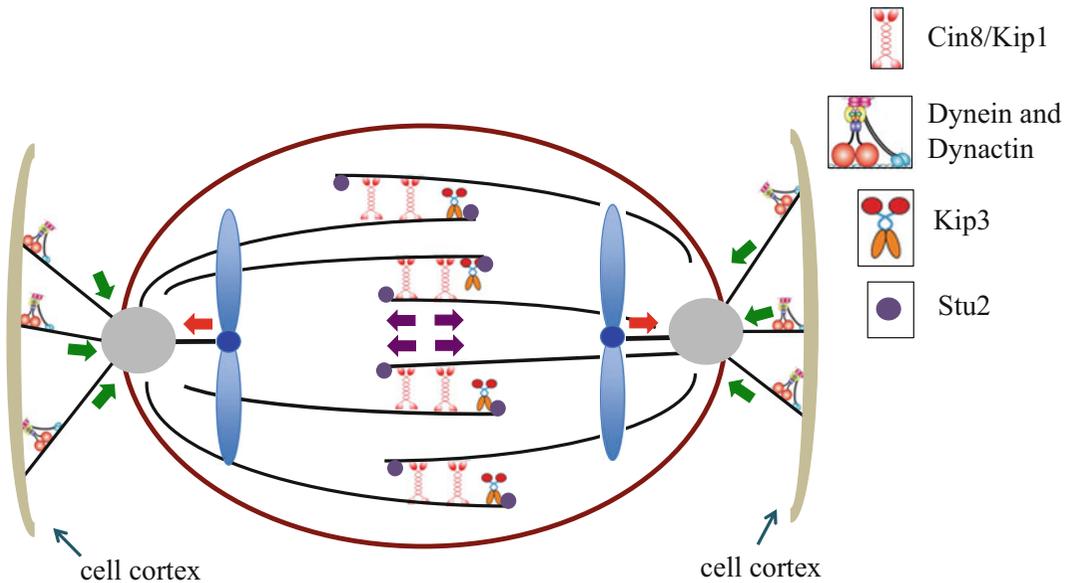
Cin8/Kip1

Dynein and Dynactin

Kip3

Stu2

cell cortex   cell cortex

Fig. 4 During anaphase B the spindle poles move away from each other thanks to pushing forces generated at the spindle midzone through motor proteins (Cin8 and Kip1) that cause the sliding of interdigitated microtubules (*purple arrows*). Then the spindle elongation is completed by pulling forces generated by motor proteins (Dynein in complex with Dynactin) associated with astral microtubules that link the spindle to the cell cortex (*green arrows*)

elongation is completed by pulling forces generated by motor proteins associated with astral microtubules that link the spindle to the cell cortex (Yeh et al. 1995) (Fig. 4).

Among other processes, the elongation of the mitotic spindle relies on a balance between phosphorylation and dephosphorylation events, that have to occur sequentially and have to be tightly coordinated to cell cycle progression. A key player in spindle elongation is Ase1, a component of the spindle midzone that regulates proper localization of almost all midzone proteins. Ase1 is required to midzone formation, to promote spindle elongation and stability in anaphase and plays also a role in cytokinesis (Pellman et al. 1995; Schuyler et al. 2003; Kotwaliwale et al. 2007; Norden et al. 2006). Ase1 is regulated by phosphorylation during the cell cycle: it is phosphorylated by Cdk1 during metaphase and this modification is required for spindle stability in anaphase (Juanes et al. 2011), while with anaphase onset Ase1 is dephosphorylated by Cdc14 in order to assemble a proper spindle midzone (Khmelinskii et al. 2007). Ase1 phosphorylation

does not change its microtubule binding activity or its ability to dimerize, instead, the different phosphorylation state of Ase1 contributes to the differential recruitment of motor proteins. In fact, in metaphase the phosphorylated Ase1 prevents the association of Cin8 with the overlapping microtubules of the spindle midzone so the microtubule sliding forces are restrained. Dephosphorylation of Ase1 upon anaphase onset is necessary and sufficient to recruit the motor protein Cin8 to the midzone thereby promoting active spindle elongation.

Besides stabilizing microtubules dynamics and spindle midzone, it has been reported recently that FEAR-activated Cdc14 plays an essential function in mitotic spindle elongation: it directly dephosphorylates and activates the motor protein Cin8 (Roccuzzo et al. 2015). The authors defined the presence of a "restriction point" for anaphase commitment that comes after the cohesin cleavage and that inhibits spindle elongation. The activity of the polo-like kinase Cdc5 and of Cdc14 are strictly required for mitotic spindle elongation, importantly the

authors demonstrate that Cin8 is a key target of Cdc14 and that its dephosphorylation is crucial to drive anaphase spindle elongation.

During anaphase B, mitotic spindle elongates thanks to the polymerization activity of Stu2 (Severin et al. 2001). The kinesin Kip3 is a multifunctional protein and recently it has been shown that the activities of depolymerase, plus-end motility and antiparallel sliding are all required to control spindle length (Rizk et al. 2014). In particular, during anaphase Kip3 counteracts Stu2 action, limits the region of antiparallel MT overlap and ensures that spindle elongation terminates when the spindle reaches the proper final length.

Chromosome passenger complexes (CPCs) are other important players in spindle dynamics. CPCs are composed of Aurora B kinase (Ipl1), Sli15, Slk19 Bir1 and Nbl1, they associate with kinetochores and help correct attachment to microtubules and, after anaphase onset, they localize to the anaphase spindle (Adams et al. 2001), this dynamic localization in time and space suggests that CPCs are involved in several steps of mitosis. Indeed, they are required to maintain spindle midzone organization (Khmelinskii and Schiebel 2008) and a CPC complex without Ipl1 also play a role in mitotic spindle elongation, as it is required for maximal spindle elongation rate (Rozelle et al. 2011). CPCs may increase outward sliding forces by activating motors or by inhibition of a brake that could be of different molecular nature (a protein or a crosslink between antiparallel interpolar MTs). Recent work demostrates that CPC positively regulates Cin8 and Kip1 thus stimulating outward sliding (Rozelle et al. 2011). In addition, CPCs are required to correctly localize Cin8 and Kip1 in the middle of the spindle likely by altering posttanslational modifications of spindle midzone proteins or through direct interaction with spindle proteins that act as spacers. These findings support the hypothesis that CPCs are important to ensure proper chromosome segregation as they control the spindle length in response to lagging chromosomes or to mitotic spindle defects. Indeed, a "no-cut" pathway has been described that is conserved from yeast to human (Norden

et al. 2006). Ipl1 and Sli15 act at the top of the cascade, they signal eventual spindle problems or unsegregated chromosomes and induce a delay in cytokinesis in order to avoid chromosome breakage due to clevage furrow ingression (Mendoza et al. 2009). In summary, the spindle midzone and CPCs coordinate spindle elongation with cytokinesis completion by inducing a spindle brake and by blocking abscission.

5   Chapter 3. Mitotic Spindle Disassembly at the End of Mitosis

At the end of anaphase, the spindle has very short kMTs and some long MTs that interdigitate at the spindle midzone, this structure must be disassembled before cytokinesis. The spindle disassembly can be divided in two processes: spindle splitting and microtubule depolymerization. Spindle splitting is caused by APC activation that leads to the degradation of Ase1 and Cin8 that crosslink MTs in the spindle midzone (Hildebrandt and Hoyt 2001; Juang et al. 1997). Microtubule depolymerization is brought about by the microtubule depolymerase Kip3 and, in parallel, by the kinase Ipl1 that has a dual role: it phosphorylates and inactivates the MT plus end binding protein Bim1 (Woodruff et al. 2010; Tirnauer et al. 1999) and phosphorylates and activates the spindle destabilizing protein She1 (Woodruff et al. 2010).

The spindle disassembly is partially regulated by the MEN, that coordinates sister chromatids separation, inactivation of mitotic CDKs and microtubule depolymerization. MEN activation leads to Cdc14-mediated dephosphorylation of Cdh1, the APC cofactor, and allows APC-mediated degradation of Cin8, Ase1 and Fin1, all involved in spindle stabilization (Woodbury and Morgan 2007; Hildebrandt and Hoyt 2001; Juang et al. 1997) and of mitotic cyclins that helps CDK inactivation.

The process of mitotic spindle disassembly must be completed in a correct way in order to preserve cell viability, indeed defects in this pathway causes lethality. Daughter cells that inherit a partially disassembled spindle are not

able to form a functional bipolar spindle and to undergo a normal mitosis (Woodruff et al. 2012). In these cells SPBs duplication and separation is normal but microtubule polymerization is impaired, so it has been proposed that complete spindle disassembly is essential for regeneration of the tubulin pool necessary for efficient spindle assembly (Woodruff et al. 2012).

6   Concluding Remarks

The mitotic spindle is essential for cell proliferation and for maintenance of cell viability due to its crucial role in segregating chromosomes at daughter cells during mitosis. The spindle is a very dynamic structure that is assembled de novo, elongated and disassembled during each round of cell division and in perfect coordination with all other cell cycle events. These processes are very complex and a lot of proteins are required to ensure their accuracy and fidelity. In the past years genetics, biochemistry, proteomics, structural biology, electron and light microscopy allowed the researchers to collect plenty of data to deepen our knowledge on mitotic spindle structure and function. However there are still several issues to be solved and challenges to be taken up, and the development of new technologies such as computational and network modelling, super-resolution and single molecule microscopy will provide us new instruments to study the fascinating field of the mitotic spindle.

Statement The authors declare that this manuscript has not been published elsewhere and it has not been submitted for publication elsewhere. All authors agree to the submission to the journal.

Compliance with Ethical Standards Conflicts of interest The authors declare no conflicts of interest.

Ethical approval This article do not contain any studies with human participants or animals performed by any of the authors.

References

Adams IR, Kilmartin JV (1999) Localization of core spindle pole body (SPB) components during SPB duplication in Saccharomyces cerevisiae. J Cell Biol 145(4):809–823

Adams RR, Carmena M, Earnshaw WC (2001) Chromosomal passengers and the (aurora) ABCs of mitosis. Trends Cell Biol 11(2):49–54

Booher RN, Deshaies RJ, Kirschner MW (1993) Properties of Saccharomyces cerevisiae wee1 and its differential regulation of p34CDC28 in response to G1 and G2 cyclins. EMBO J 12(9):3417–3426

Breitkreutz A, Choi H, Sharom JR, Boucher L, Neduva V, Larsen B, Lin ZY, Breitkreutz BJ, Stark C, Liu G, Ahn J, Dewar-Darch D, Reguly T, Tang X, Almeida R, Qin ZS, Pawson T, Gingras AC, Nesvizhskii AI, Tyers M (2010) A global protein kinase and phosphatase interaction network in yeast. Science 328(5981):1043–1046

Bullitt E, Rout M, Kilmartin J, Akey C (1997) The yeast spindle pole body is assembled around a central crystal of Spc42p. Cell 89:1077–1086

Carvalho P, Tirnauer JS, Pellman D (2003) Surfing on microtubule ends. Trends Cell Biol 13:229–237

Carvalho P, Gupta ML Jr, Hoyt MA, Pellman D (2004) Cell cycle control of kinesin-mediated transport of Bik1 (CLIP-170) regulates microtubule stability and dynein activation. Dev Cell 6:815–829

Chiroli E, Rancati G, Catusi I, Lucchini G, Piatti S (2009) Cdc14 inhibition by the spindle assembly checkpoint prevents unscheduled centrosome separation in budding yeast. Mol Biol Cell 20(10):2626–2637

Clarke DJ, Bachant J (2008) Kinetochore structure and spindle assembly checkpoint signaling in the budding yeast, Saccharomyces cerevisiae. Front Biosci 13:6787–6819

Crasta K, Huang P, Morgan G, Winey M, Surana U (2006) Cdk1 regulates centrosome separation by restraining proteolysis of microtubule-associated proteins. EMBO J 25(11):2551–2563

Crasta K, Lim HH, Giddings TH Jr, Winey M, Surana U (2008) Inactivation of Cdh1 by synergistic action of Cdk1 and polo kinase is necessary for proper assembly of the mitotic spindle. Nat Cell Biol 10(6):665–675

De Simone A, Nédélec F, Gönczy P (2016) Dynein transmits polarized actomyosin cortical flows to promote centrosome separation. Cell Rep 14 (9):2250–2262

Gheber L, Kuo SC, Hoyt MA (1999) Motile properties of the kinesin-related Cin8p spindle motor extracted from Saccharomyces cerevisiae cells. J Biol Chem 274:9564–9572

Gupta ML Jr, Carvalho P, Roof DM, Pellman D (2006) Plus end-specific depolymerase activity of Kip3, a kinesin-8 protein, explains its role in positioning the yeast mitotic spindle. Nat Cell Biol 8:913–923

Haase SB, Winey M, Reed SI (2001) Multi-step control of spindle pole body duplication by cyclin-dependent-kinase. Nat Cell Biol 3:38–42

Higuchi T, Uhlmann F (2005) Stabilization of microtubule dynamics at anaphase onset promotes chromosome segregation. Nature 433(7022):171–176

Hildebrandt ER, Hoyt MA (2000) Mitotic motors in Saccharomyces cerevisiae. Biochim Biophys Acta 1496:99–116

Hildebrandt ER, Hoyt MA (2001) Cell cycle-dependent degradation of the Saccharomyces cerevisiae spindle motor Cin8p requires APC(Cdh1) and a bipartite destruction sequence. Mol Biol Cell 12(11):3402–3416

Hwang E, Kusch J, Barral Y, Huffaker TC (2003) Spindle orientation in Saccharomyces cerevisiae depends on the transport of microtubule ends along polarized actin cables. J Cell Biol 161:483–488

Jacobs CW, Adams AE, Szaniszlo PJ, Pringle JR (1988) Functions of microtubules in the Saccharomyces cerevisiae cell cycle. J Cell Biol 107:1409–1426

Jaspersen SL, Winey M (2004) The budding yeast spindle pole body: structure, duplication, and function. Annu Rev Cell Dev Biol 20:1–28

Juanes MA, ten Hoopen R, Segal M (2011) Ase1p phosphorylation by cyclin-dependent kinase promotes correct spindle assembly in S. cerevisiae. Cell Cycle 10(12):1988–1997

Juang YL, Huang J, Peters JM, McLaughlin ME, Tai CY, Pellman D (1997) APC-mediated proteolysis of Ase1 and the morphogenesis of the mitotic spindle. Science 275:1311–1314

Kahana JA, Schlenstedt G, Evanchuk DM, Geiser JR, Hoyt MA, Silver PA (1998) The yeast dynactin complex is involved in partitioning the mitotic spindle between mother and daughter cells during anaphase B. Mol Biol Cell 9(7):1741–1756

Kellogg DR (2003) Wee1-dependent mechanisms required for coordination of cell growth and cell division. J Cell Sci 116(Pt 24):4883–4890

Khmelinskii A, Schiebel E (2008) Assembling the spindle midzone in the right place at the right time. Cell Cycle 7:283–286

Khmelinskii A, Lawrence C, Roostalu J, Schiebel E (2007) Cdc14-regulated midzone assembly controls anaphase B. J Cell Biol 177:981–993

Khmelinskii A, Roostalu J, Roque H, Antony C, Schiebel E (2009) Phosphorylation-dependent protein interactions at the spindle midzone mediate cell cycle regulation of spindle elongation. Dev Cell 17:244–256

Khodjakov A, La Terra S, Chang F (2004) Laser microsurgery in fission yeast; role of the mitotic spindle midzone in anaphase B. Curr Biol 14(15):1330–1340

King K, Kang H, Jin M, Lew DJ (2013) Feedback control of Swe1p degradation in the yeast morphogenesis checkpoint. Mol Biol Cell 24(7):914–922

Kosco KA, Pearson CG, Maddox PS, Wang PJ, Adams IR, Salmon ED, Bloom K, Huffaker TC (2001) Control of microtubule dynamics by Stu2p is essential for spindle orientation and metaphase chromosome alignment in yeast. Mol Biol Cell 12(9):2870–2880

Kotwaliwale CV, Frei SB, Stern BM, Biggins S (2007) A pathway containing the Ipl1/aurora protein kinase and the spindle midzone protein Ase1 regulates yeast spindle assembly. Dev Cell 13(3):433–445

Lee WL, Kaiser MA, Cooper JA (2005) The offloading model for dynein function: differential function of motor subunits. J Cell Biol 168:201–207

Maddox P, Straight A, Coughlin P, Mitchison TJ, Salmon ED (2003) Direct observation of microtubule dynamics at kinetochores in Xenopus extract spindles implications for spindle mechanics. J Cell Biol 162(3):377–382

Maiato H, Khodjakov A, Rieder CL (2005) Drosophila CLASP is required for the incorporation of microtubule subunits into fluxing kinetochore fibres. Nat Cell Biol 7:42–47

Martinez JS, Jeong DE, Choi E, Billings BM, Hall MC (2006) Acm1 is a negative regulator of the CDH1-dependent anaphase-promoting complex/cyclosome in budding yeast. Mol Cell Biol 26(24):9162–9176

McMillan JN, Theesfeld CL, Harrison JC, Bardes ES, Lew DJ (2002) Determinants of Swe1p degradation in Saccharomyces cerevisiae. Mol Biol Cell 13(10):3560–3575

Mendoza M, Norden C, Durrer K, Rauter H, Uhlmann F, Barral Y (2009) A mechanism for chromosome segregation sensing by the NoCut checkpoint. Nat Cell Biol 11:477–483

Molk JN, Bloom K (2006) Microtubule dynamics in the budding yeast mating pathway. J Cell Sci 119:3485–3490

Moore JK, Stuchell-Brereton MD, Cooper JA (2009) Function of dynein in budding yeast: mitotic spindle positioning in a polarized cell. Cell Motil Cytoskeleton 66:546–555

Norden C, Mendoza M, Dobbelaere J, Kotwaliwale CV, Biggins S, Barral Y (2006) The NoCut pathway links completion of cytokinesis to spindle midzone function to prevent chromosome breakage. Cell 125(1):85–98

O'Toole E, Winey M, McIntosh JR (1999) High-voltage electron tomography of spindle pole bodies and early mitotic spindles in the yeast Saccharomyces cerevisiae. Mol Biol Cell 10:2017–2031

Pasqualone D, Huffaker TC (1994) STU1, a suppressor of a beta-tubulin mutation, encodes a novel and essential component of the yeast mitotic spindle. J Cell Biol 127(6 Pt 2):1973–1984

Pearson CG, Bloom K (2004) Dynamic microtubules lead the way for spindle positioning. Nat Rev Mol Cell Biol 5(6):481–492

Pellman D, Bagget M, Tu YH, Fink GR, Tu H (1995) Two microtubule-associated proteins required for anaphase spindle movement in Saccharomyces cerevisiae. J Cell Biol 130(6):1373–1385

Pereira G, Schiebel E (2003) Separase regulates INCENP-Aurora B anaphase spindle function through Cdc14. Science 302:2120–2124

Pereira G, Tanaka TU, Nasmyth K, Schiebel E (2001) Modes of spindle pole body inheritance and segregation of the Bfa1p–Bub2p checkpoint protein complex. EMBO J 20(22):6359–6370

Raspelli E, Cassani C, Lucchini G, Fraschini R (2011) Budding yeast Dma1 and Dma2 participate in regulation of Swe1 levels and localization. Mol Biol Cell 22 (13):2185–2197

Raspelli E, Cassani C, Chiroli E, Fraschini R (2015) Budding yeast Swe1 is involved in the control of mitotic spindle elongation and is regulated by Cdc14 phosphatase during mitosis. J Biol Chem 290(1):1–12

Rizk RS, DiScipio KA, Proudfoot KG, Gupta MR Jr (2014) The kinesin-8 Kip3 scales anaphase spindle length by suppression of midzone microtubule polymerization. J Cell Biol 204(6):965–975

Roccuzzo M, Visintin C, Tili F, Visintin R (2015) FEAR-mediated activation of Cdc14 is the limiting step for spindle elongation and anaphase progression. Nat Cell Biol 17(3):251–261

Rogers GC, Rogers SL, Sharp DJ (2005) Spindle microtubules in flux. J Cell Sci 118(Pt 6):1105–1116

Roof DM, Meluh PB, Rose MD (1992) Kinesin-related proteins required for assembly of the mitotic spindle. J Cell Biol 118(1):95–108

Rozelle DK, Hansen SD, Kaplan KB (2011) Chromosome passenger complexes control anaphase duration and spindle elongation via a kinesin-5 brake. J Cell Biol 193:285–294

Saunders WS, Hoyt MA (1992) Kinesin-related proteins required for structural integrity of the mitotic spindle. Cell 70(3):451–458

Schuyler SC, Liu JY, Pellman D (2003) The molecular function of Ase1p: evidence for a MAP-dependent midzone-specific spindle matrix. J Cell Biol 160 (4):517–528

Segal M, Clarke DJ, Maddox P, Salmon ED, Bloom K, Reed SI (2000) Coordinated spindle assembly and orientation requires Clb5-dependent kinase in budding yeast. J Cell Biol 148:441–451

Severin F, Habermann B, Huffaker T, Hyman T (2001) Stu2 promotes mitotic spindle elongation in anaphase. J Cell Biol 153(2):435–442

Stegmeier F, Amon A (2004) Closing mitosis: the functions of the Cdc14 phosphatase and its regulation. Annu Rev Genet 38:203–232

Stegmeier F, Visintin R, Amon A (2002) Separase, polo kinase, the kinetochore protein Slk19, and Spo12 function in a network that controls Cdc14 localization during early anaphase. Cell 108(2):207–220

Tirnauer JS, O'Toole E, Berrueta L, Bierer BE, Pellman D (1999) Yeast Bim1p promotes the G1-specific dynamics of microtubules. J Cell Biol 145:993–1007

Vaisberg EA, Koonce MP, McIntosh JR (1993) Cytoplasmic dynein plays a role in mammalian mitotic spindle formation. J Cell Biol 123:849–858

Varga V, Helenius J, Tanaka K, Hyman AA, Tanaka TU, Howard J (2006) Yeast kinesin-8 depolymerizes microtubules in a length-dependent manner. Nat Cell Biol 8(9):957–962

Wang X, Tsai JW, Imai JH, Lian WN, Vallee RB, Shi SH (2009) Asymmetric centrosome inheritance maintains neural progenitors in the neocortex. Nature 461 (7266):947–955

Wargacki MM, Tay JC, Muller EG, Asbury CL, Davis TN (2010) Kip3, the yeast kinesin-8, is required for clustering of kinetochores at metaphase. Cell Cycle 9 (13):2581–2588

Waters JC, Mitchison TJ, Rieder CL, Salmon ED (1996) The kinetochore microtubule minus-end disassembly associated with poleward flux produces a force that can do work. Mol Biol Cell 7:1547–1558

Winey M, Bloom K (2012) Mitotic spindle form and function. Genetics 190(4):1197–1224

Winey M, Mamay CL, Toole ETO, Mastronarde DN, Giddings TH Jr, McDonald KL, McIntosh JR (1995) Three-dimensional ultrastructural analysis of the Saccharomyces cerevisiae mitotic spindle. J Cell Biol 129:1601–1615

Woodbury EL, Morgan DO (2007) Cdk and APC activities limit the spindle-stabilizing function of Fin1 to anaphase. Nat Cell Biol 9(1):106–112

Woodruff JB, Drubin DG, Barnes G (2010) Mitotic spindle disassembly occurs via distinct subprocesses driven by the anaphase-promoting complex, Aurora B kinase, and kinesin-8. J Cell Biol 191:795–808

Woodruff JB, Drubin DG, Barnes G (2012) Spindle assembly requires complete disassembly of spindle remnants from the previous cell cycle. Mol Biol Cell 23(2):258–267

Yeh E, Skibbens RV, Cheng JW, Salmon ED, Bloom K (1995) Spindle dynamics and cell cycle regulation of dynein in the budding yeast, Saccharomyces cerevisiae. J Cell Biol 130(3):687–700

Yin H, You L, Pasqualone D, Kopski KM, Huffaker TC (2002) Stu1p is physically associated with beta-tubulin and is required for structural integrity of the mitotic spindle. Mol Biol Cell 13:1881–1892

Secreted Phospholipase A2 Type IIA (sPLA2-IIA) Activates Integrins in an Allosteric Manner

Yoshikazu Takada and Masaaki Fujita

Abstract

Secreted phospholipase A2 type IIA (sPLA2-IIA) is a well-established pro-inflammatory protein and has been a major target for drug discovery. However, the mechanism of its signaling action has not been fully understood. We previously found that sPLA2-IIA binds to integrins αvβ3 and α4β1 in human and that this interaction plays a role in sPLA2-IIA's signaling action. Our recent studies found that sPLA2-IIA activates integrins in an allosteric manner through direct binding to a newly identified binding site of integrins (site 2), which is distinct from the classical RGD-binding site (site 1). The sPLA2-IIA-induced integrin activation may be related to the signaling action of sPLA2-IIA. Since sPLA2-IIA is present in normal human tears in addition to rheumatoid synovial fluid at high concentrations the sPLA2-IIA-mediated integrin activation on leukocytes may be involved in immune responses in normal and pathological conditions.

Keywords

Secreted phospholipase A2 type IIA (sPLA2-IIA) • Integrin activation • Allosteric site • Docking simulation

Y. Takada (✉)
Department of Dermatology, Biochemistry and
Molecular Medicine, UC Davis School of Medicine,
Research III Suite 3300, 4645 Second Avenue,
Sacramento, CA 95817, USA

The PhD Program for Translational Medicine, College of
Medical Science and Technology, Taipei Medical
University, 250 Wu-Hsing Street, Taipei 11031, Taiwan,
Republic of China
e-mail: ytakada@ucdavis.edu

M. Fujita
Department of Clinical Immunology and Rheumatology,
The Tazuke-Kofukai Medical Research Institute, Kitano
Hospital, 2-4-20 Ohgimachi, Kita-ku,
Osaka 530-8480, Japan
e-mail: ms-fujita@kitano-hp.or.jp

Abbreviations

FKN-CD	the chemokine domain of fractalkine
IL	interleukin
NF-kB	nuclear factor kappa-light-chain-enhancer of activated B cells
RGD	Arg-Gly-Asp
sPLA2-IIA	Secreted phospholipase A2 type IIA
TNF	Tumor necrosis factor.

1 What Is sPLA2-IIA?

The phospholipase A2 (PLA2) family is a group of intracellular and secreted enzymes that hydrolyzes the *sn*-2 ester bond in the glyceroacyl phospholipids present in lipoproteins and cell membranes to form nonesterified fatty acids and lysophospholipids. These products act as intracellular second messengers or are further metabolized into potent mediators of a broad range of cellular processes, including inflammation, apoptosis, and atherogenesis (Tatulian 2001). The mammalian secreted PLA2 isoforms are comprised of the groups named IB, IIA, IIC, IID, IIE, IIF, V, X, and XII (Six and Dennis 2000; Gelb et al. 2000). All secreted PLA2 isoforms have a Ca^{2+}-dependent catalytic mechanism, a low molecular mass (13–16 kDa), several disulfide linkages, and a well-conserved overall three-dimensional structure (Six and Dennis 2000; Gelb et al. 1999; Valentin and Lambeau 2000). Secreted PLA2 type IIA (sPLA2-IIA) was first isolated and purified from rheumatoid synovial fluid (Vadas et al. 1985). sPLA2-IIA is an acute phase reactant and its plasma concentrations markedly increase in diseases that involve systemic inflammation such as rheumatoid arthritis, cardiovascular disease (up to 1000-fold and >1 μg/ml) and sepsis. Inflammatory cytokines such as interleukin (IL)-6, tumor necrosis factor (TNF)-α, and IL-1β induce synthesis and release of sPLA2-IIA in arterial smooth muscle cells and hepatocytes, which are the major sources of the plasma sPLA2-IIA in these systemic inflammatory conditions (Jaross et al. 2002; Niessen et al. 2003). sPLA2-IIA$^{+/+}$ mice have much more severe symptoms of rheumatoid arthritis than in sPLA2-IIA$^{-/-}$ mice, indicating that sPLA2-IIA is clearly involved in rheumatoid arthritis (Boilard et al. 2010). In addition to being a pro-inflammatory protein, sPLA2-IIA expression is elevated in neoplastic prostatic tissue (Jiang et al. 2002) and dysregulation of sPLA2-IIA may play a role in prostatic carcinogenesis (Dong et al. 2006), and is a potential therapeutic target in prostate cancer (Sved et al. 2004). Also, lung cancer cells secrete sPLA2-IIA, and plasma sPLA2-IIA can potentially serve as lung cancer biomarker and sPLA2IIA is a potential therapeutic target against lung cancer (Dong et al. 2014).

2 sPLA2-IIA Mediates Its Signaling Actions Through Receptor Binding, Not Catalytic Activity

It is expected that these biological actions are mediated by sPLA2-IIA's catalytic activity. However, some biological effects associated with sPLA2-IIA are independent of its catalytic function (Tada et al. 1998). Catalytically inactive sPLA2-IIA mutants retain the ability to enhance cyclooxygenase-2 expression in connective tissue mast cells (Tada et al. 1998). Also inactivation of sPLA2-IIA by bromophenacyl bromide does not affect the ability of sPLA2-IIA to induce secretion of β-glucuronidase, IL-6, and IL-8 from human eosinophils (Triggiani et al. 2003). It has thus been proposed that sPLA2-IIA action is mediated through interaction with specific receptors. Indeed the enzyme binds to a high affinity receptor of 180 kDa present on rabbit skeletal muscle (Lambeau et al. 1994). This so-called M (muscle)-type receptor belongs to the superfamily of C-type lectins and mediates some of the physiological effects of mammalian sPLA2-IIA, and binding of sPLA2-IIA to this receptor induces internalization of sPLA2-IIA (Nicolas et al. 1995). However, the interaction

between sPLA2-IIA and the M-type receptor is species-specific, and human sPLA2-IIA binds to the human or mouse M-type receptor very weakly (Cupillard et al. 1999). Thus, sPLA2-IIA receptors in human have not been established. Mammalian sPLA2-IIAs bind to heparan sulfate proteoglycans like glypican-1 (Murakami et al. 1999) and decorin in apoptotic human T cells (Sartipy et al. 2000). The binding of sPLA2-IIA to heparan sulfate proteoglycans has been implicated in the release of arachidonic acid from apoptotic T cells (Boilard et al. 2003), but it is unclear whether this process plays a role in other situations.

3 Integrins Are Receptors for sPLA2-IIA

Our previous studies found that sPLA2-IIA binds to integrins αvβ3 and α4β1 (Saegusa et al. 2008). Integrins are a family of cell adhesion receptors that recognize extracellular matrix ligands (e.g., collagen and fibronectin), cell surface ligands (e.g., intracellular adhesion molecules, ICAMs; vascular cell adhesion molecule-1, VCAM-1), and small soluble ligands (e.g., fibroblast growth factor and insulin-like growth factor) (Hynes 2002; Takada et al. 2007). Integrins are transmembrane heterodimers, and at least 18 α and 8 β subunits are known (Takada et al. 2007). Integrins transduce signals to the cell upon ligand binding (Hynes 2002). We found that sPLA2-IIA specifically binds to integrin αvβ3 and α4β1 and induces proliferative signals in an integrin-dependent manner (Saegusa et al. 2008). The integrin-binding site does not include the catalytic center or the M-type receptor-binding site. Consistently, WT and the catalytically inactive mutant (the H47Q mutant) of sPLA2-IIA induce intracellular signals in monocytic cells. In contrast, the integrin-binding defective mutant (the R74E/R100E mutant) does not (Saegusa et al. 2008). These results suggest that integrins may serve as receptors for sPLA2-IIA and mediate signaling action of sPLA2-IIA in human.

4 Small Molecule Antagonists for sPLA2-IIA-Integrin Interaction

The findings suggest that sPLA2-IIA binding to integrins is a therapeutic target. We looked for small compounds bind to the integrin-binding site of sPLA2-IIA (around R74 and R100). We selected compounds that bind only to WT sPLA2-IIA (hopefully only to the integrin-binding site) by screening one-bead-one compound (OBOC) library using the R74E/R100E mutant and WT sPLA2-IIA (Ye et al. 2013). We obtained several compounds including compound 21 (**Cmpd21**) that suppress αvβ3-mediated cell adhesion and migration (Ye et al. 2013). These findings indicate direct binding of sPLA2-IIA to integrins is indeed critical for signaling actions of sPLA2-IIA, which is consistent with our model (Saegusa et al. 2008).

5 Allosteric Activation of Integrins (Historic Overview)

It has been well established that integrin activation is mediated by signaling from inside the cell (inside-out signaling), and that integrin activation is associated with global conformational changes of the integrin molecule (Zhu et al. 2013; Xiong et al. 2009). Our recent studies found, however, that integrins can be activated by chemokine without its receptor and this is not consistent with the inside-out signaling model.

It has previously been reported that the binding of an RGD-mimetic peptide induces changes in the tertiary structure of αvβ3 (Xiong et al. 2002) and αIIbβ3 (Xiao et al. 2004) in the β3 I-like domain. RGD or ligand-mimetic peptides activate purified, non-activated αIIbβ3 (Du et al. 1991) and αvβ3 (Legler et al. 2001). This process does not require inside-out signal transduction and it appears that RGD or ligand-mimetic peptide triggers conformational changes that lead to full activation of integrins. These findings suggest that these peptides enhance integrin affinity by conformational changes in

the headpiece possibly through additional ligand-binding sites in the integrin (Du et al. 1991). A previous study suggests that there are two RGD-binding sites in integrin αIIbβ3, and that one binding site acts as an allosteric site based on binding kinetic studies (Hu et al. 1999). Also, another study suggests that two distinct cyclic RGD-mimetic peptides can simultaneously bind to distinct sites in αIIbβ3, and the estimated distance between two ligand-binding site is about 6.1 ± 0.5 nm (Cierniewski et al. 1999). The possible allosteric ligand-binding site has not been pursued probably because the αvβ3 structure (ligand occupied, open) contains only one RGD-binding site (Xiong et al. 2002).

6 Allosteric Activation of Integrins Through Binding to a Newly Identified Ligand-Binding Site (Site 2)

We discovered that the chemokine domain of membrane-bound chemokine fractalkine (FKN-CD) directly binds to several integrins and this interaction is critical for fractalkine/CX3CR1 signaling (Fujita et al. 2012a). FKN-CD induces ternary complex formation (integrin-FKN-CD-CX3CR1) on the cell surface, suggesting that integrins act as co-receptor for FKN-CD in FKN/CX3CR1 signaling (Fujita et al. 2012a).

Notably we discovered that FKN-CD markedly enhances ligand (e.g., a fibrinogen fragment, γC399tr., Yokoyama et al. 1999, 2000) binding to purified soluble integrin αvβ3 in cell-free conditions (Fujita et al. 2014). Since no chemokine receptor is present, we suspected that this activation requires direct binding to integrins. Only a single binding site for ligand (e.g., RGD) has been identified in integrin crystal structures (e.g., 1L5G.pdb for ligand-bound αvβ3). This does not, however, explain the mechanism of FKN-CD-induced integrin activation in cell-free conditions. If FKN-CD and the fibrinogen fragment bind to the same binding site, FKN-CD is expected to inhibit the binding of the fibrinogen fragment, but do not enhance

it. We suspected that the CX3CR1-independent integrin activation by FKN-CD is mediated by the binding of FKN-CD to another binding site that is not known. We also suspected that the second ligand-binding site (possible allosteric site) may be open in inactive integrins, but is closed in activated form. To predict the location of such a binding site, we performed docking simulation using inactive αvβ3 as a target (1JV2.pdb) (Fig. 1). Notably, the simulation predicted a potential FKN-CD-binding site in the inactive αvβ3 (site 2) that is distinct from the classical RGD-binding site (site 1) (Fujita et al. 2014). Site 2 is located at a crevice between αv and β3 on the opposite side of site 1 in the αvβ3 headpiece. A 21-mer peptide from site 2 of αvβ3 (S2-β3 peptide, residues 267–286 of β3, as glutathion-S-transferase fusion protein) specifically binds to FKN-CD and effectively suppresses integrin activation by FKN-CD, suggesting that FKN-CD actually binds to site 2. We thus propose a model, in which FKN-CD binding to site 2 induces activation of site 1 though conformational changes (in an allosteric mechanism) (Fujita et al. 2014).

In our docking model the distance between site 1 and site 2 is about 6 nm. Thus, the position of site 2 is consistent with the previous report. Based on previous studies it is likely that the newly identified site 2 has ligand specificity that overlaps with that of site 1, interacts with integrin ligands other than FKN-CD (e.g., RGD), and is potentially involved in integrin regulation in an allosteric mechanism. It is reasonable to assume that FKN-CD binding to site 2 induces conformational changes in integrins. We suspect that other proteins also bind to site 2 and affect integrin functions. Also it is likely that integrins other than αvβ3, α4β1, and α5β1 are activated by site 2-mediated mechanism.

7 sPLA2-IIA Activates Soluble Integrin αvβ3 in Cell-Free Conditions

We studied if sPLA2-IIA enhances the binding of recombinant soluble αvβ3 to the fibrinogen

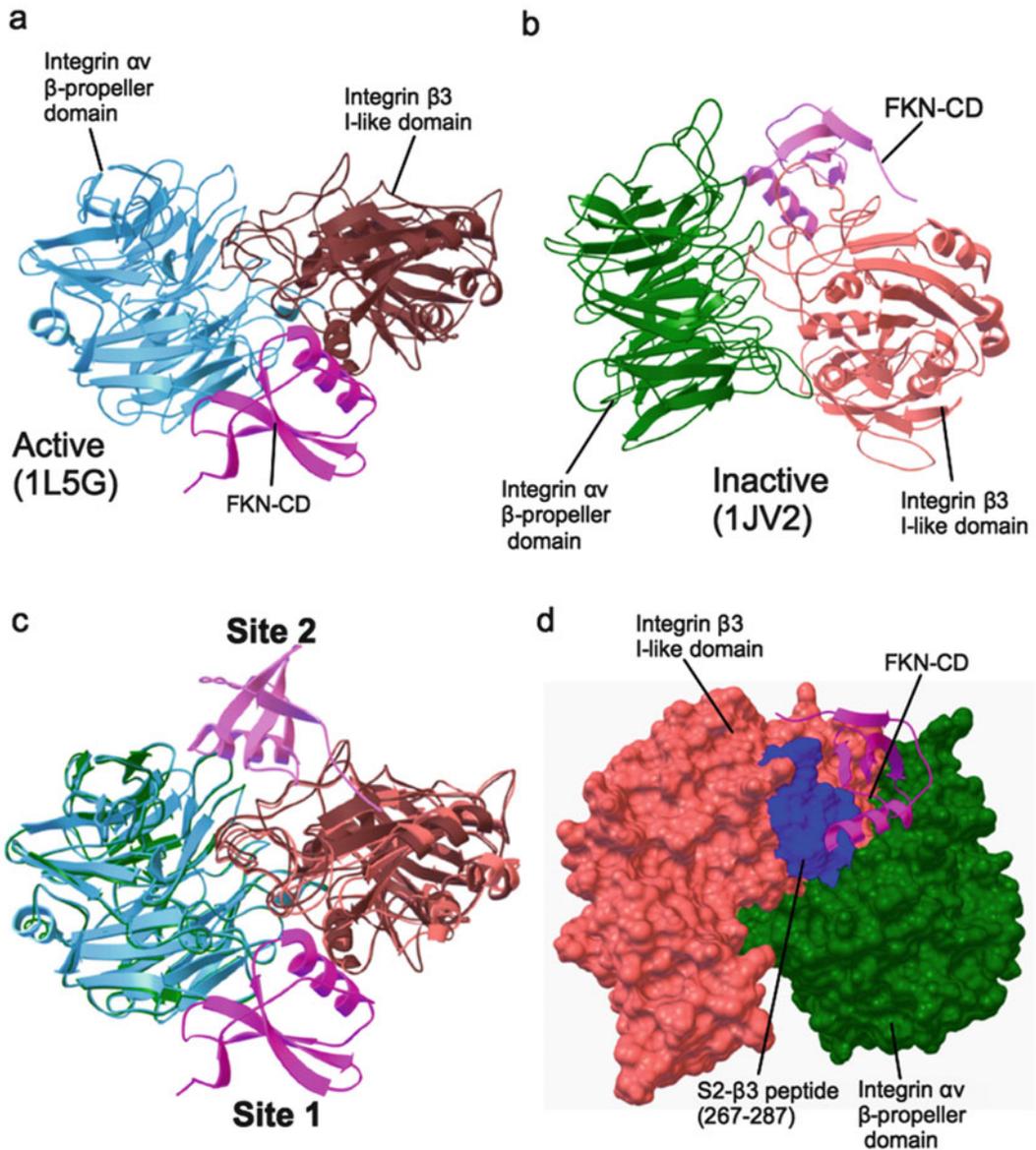**Fig. 1 Prediction of an allosteric ligand-binding site (site 2).** (**a**) A docking model of FKN-CD-integrin αvβ3 (active) interaction (Fujita et al. 2012a). The headpiece of ligand-bound form of integrin αvβ3 (PDB code 1L5G) was used as a target. The model predicts that FKN-CD (PDB code 1F2L, *red*) binds to the classical RGD-binding site of the integrin αvβ3 headpiece (site 1). (**b**) A docking model of FKN-CD-integrin αvβ3 (inactive) interaction. The headpiece of an inactive form of integrin αvβ3 (PDB code 1JV2) was used as a target. The model predicts the position of the second FKN-CD-binding site (site 2). (**c**) Superposition of two models shows that the positions of two predicted FKN-CD binding sites are distinct. (**d**) Position of the β3 peptide (267–287, *blue*) in site 2 (S2-β3)

fragment in cell-free conditions. We immobilized the fibrinogen fragment to wells of microtiter plates and measured the binding of soluble αvβ3 to the fibrinogen fragment in the presence of sPLA2-IIA. To keep soluble integrin inactive we included 1 mM Ca^{2+} in the assay (Fig. 2). WT sPLA2-IIA enhanced the binding of the fibrinogen fragment to αvβ3 in a
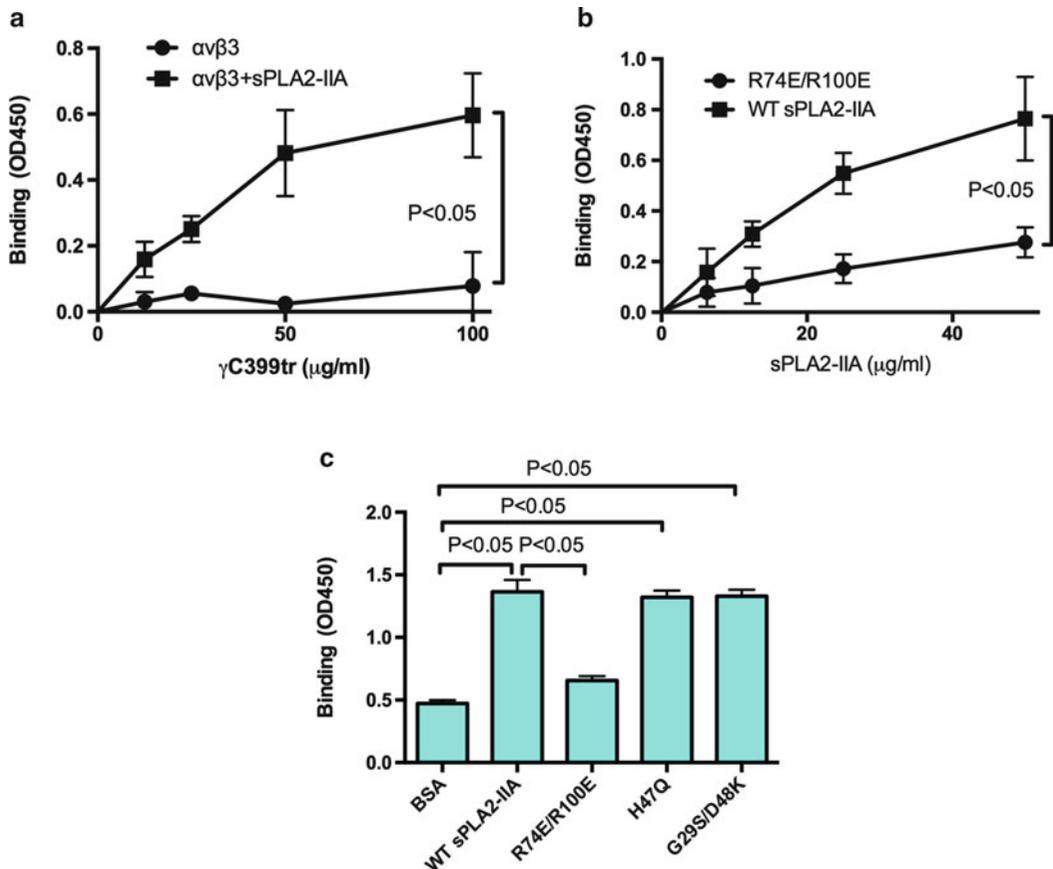
Fig. 2 sPLA2-IIA activates αvβ3 integrin in cell-free conditions (through direct integrin binding). We studied whether sPLA2-IIA directly activates integrins without inside-out signaling using recombinant soluble αvβ3 extracellular domain (no transmembrane and cytoplasmic domains). Wells of 96-well microtiter plates were coated with the fibrinogen fragment and incubated with soluble αvβ3. R74E/R100E, defective in integrin binding; H74Q, catalytically inactive; G29S/D48K, defective in the binding to M-type receptor. (**a**) activation of soluble αvβ3 by sPLA2-IIA as a function of the fibrinogen fragment concentration. (**b**) activation of soluble αvβ3 by sPLA2-IIA as a function of sPLA2-IIA concentration. (**c**) the effects of sPLA2-IIA mutations on integrin αvβ3 activation. Binding of soluble αvβ3 to the immobilized fibrinogen fragment in the presence or absence of WT sPLA2-IIA was performed

concentration-dependent manner. In contrast to WT sPLA2-IIA, the R74E/R100E mutant (integrin-binding defective) (Saegusa et al. 2008) was defective in this function. H47Q (catalytically inactive) and G29S/D48K (M-type receptor-binding defective) mutants behaved like WT sPLA2-IIA. These findings suggest that sPLA2-IIA activates αvβ3 in cell-free conditions and this activation requires the integrin-binding site of sPLA2-IIA but does not require catalytic activity or receptor binding.

We established that sPLA2-IIA activates integrins in an allosteric manner as in the case of FKN-CD. This is through direct binding of sPLA2-IIA to a newly identified binding site (site 2), which is distinct from the classical RGD-binding site (site 1) (Fujita et al. 2014). We studied if sPLA2-IIA activates integrins on the cell surface by measuring the binding of the fibrinogen fragment to cells using flow cytometry. WT sPLA2-IIA activated αvβ3 on the cell surface. The effects of sPLA2-IIA mutations on sPLA2-IIA-induced αvβ3

activation were similar to those in soluble αvβ3. These findings suggest that sPLA2-IIA activates αvβ3 on the cell surface and soluble αvβ3, and that the sPLA2-IIA-induced αvβ3 activation is not cell-type specific.

8 Docking Simulation Predicts that sPLA2-IIA Binds to Site 2 in an Inactive Form of αvβ3

We studied if sPLA2-IIA-induced activation of αvβ3 involves the binding of sPLA2-IIA to site 2 of αvβ3. Docking simulation of the interaction between sPLA2-IIA and the closed form of αvβ3 (PDB code 1JV2) predicts that sPLA2-IIA binds to site 2 with high affinity (docking energy -22.1 kcal/mol) (Fig. 3), as in the case of FKN-CD (Fujita et al. 2014). The RGD peptide binds to site 1 in the activated αvβ3. The docking model predicts that Arg74 and Arg100, which are critical for integrin binding to site 1, are within the sPLA2-IIA/αvβ3 interface at site 2, suggesting that the integrin binding interface in sPLA2-IIA at site 2 overlaps with that of site 1. This predicts that sPLA2-IIA may activate integrins through direct binding to site 2, and that the R74E/R100E mutant may be defective in this function.

9 sPLA2-IIA Directly Binds to a Peptide Derived from Site 2 of Integrin β1

We already described peptide sequences (e.g., residues 256–288 of β3, S2-β3 peptide) from site 2 that directly interacts with FKN-CD (Fujita et al. 2014). The S2-β3 peptide suppresses FKN-CD-mediated integrin activation, but control scrambled S2-β3 peptide does not (Fujita et al. 2014). We studied if S2-β3 peptide binds to sPLA2-IIA. It was expected that site 2-derived peptides bind to sPLA2-IIA, because the amino acid residues in site 2-derived peptides are located within the integrin-binding interface of sPLA2-IIA in the docking model. Interestingly, site 2 peptides from β1 (S2-β1 peptide) bound better to sPLA2-IIA in a concentration-

dependent manner than S2-β3 peptide. Control glutathion-S-transferase or scrambled S2-β3 peptide did not bind to sPLA2-IIA. S2-β1 peptide suppressed sPLA2-IIA-mediated αvβ3 activation in U937, K562 erythroleukemia or Chinese hamster ovary (CHO) cells that express recombinant αvβ3 (αvβ3-K562 and β3-CHO cells, respectively), while control glutathion-S-transferase or S2-β3scr peptide did not. These findings suggest that sPLA2-IIA binds specifically to site 2 and that the binding of sPLA2-IIA to site 2 is critical for sPLA2-IIA-mediated αvβ3 activation.

10 sPLA2-IIA Activates α4β1 and α5β1 in a Site 2-Dependent Manner

sPLA2-IIA-induced integrin activation may be integrin-specific. We found that sPLA2-IIA enhanced the ligand binding to α4β1 or α5β1, indicating that sPLA2-IIA-mediated integrin activation is not limited to αvβ3. WT sPLA2-IIA markedly increased the ligand binding to α4β1 or α5β1, while R74E/R100E did not. S2-β1 peptide suppressed the ligand binding to α4β1 or α5β1 increased by sPLA2-IIA, while control glutathion-S-transferase or scrambled peptide did not. These results suggest that sPLA2-IIA activates integrin α4β1 or α5β1 through direct binding to site 2 in a manner similar to that of αvβ3. This observation is important since α4β1 is expressed in immune competent cells.

11 Evidence that There Are Two Ligand-Binding Sites: sPLA2-IIA Induces Integrin Activation in a Biphasic Manner

If sPLA2-IIA binds to site 1 (Saegusa et al. 2008) and site 2 (Fujita et al. 2015), it is predicted that sPLA2-IIA competes with ligands for binding to site 1. To address this question, we determined the effect of sPLA2-IIA as a function of sPLA2-IIA concentrations up to 500 μg/ml. The binding of site 1 specific ligand (H120) to α4β1 on CHO
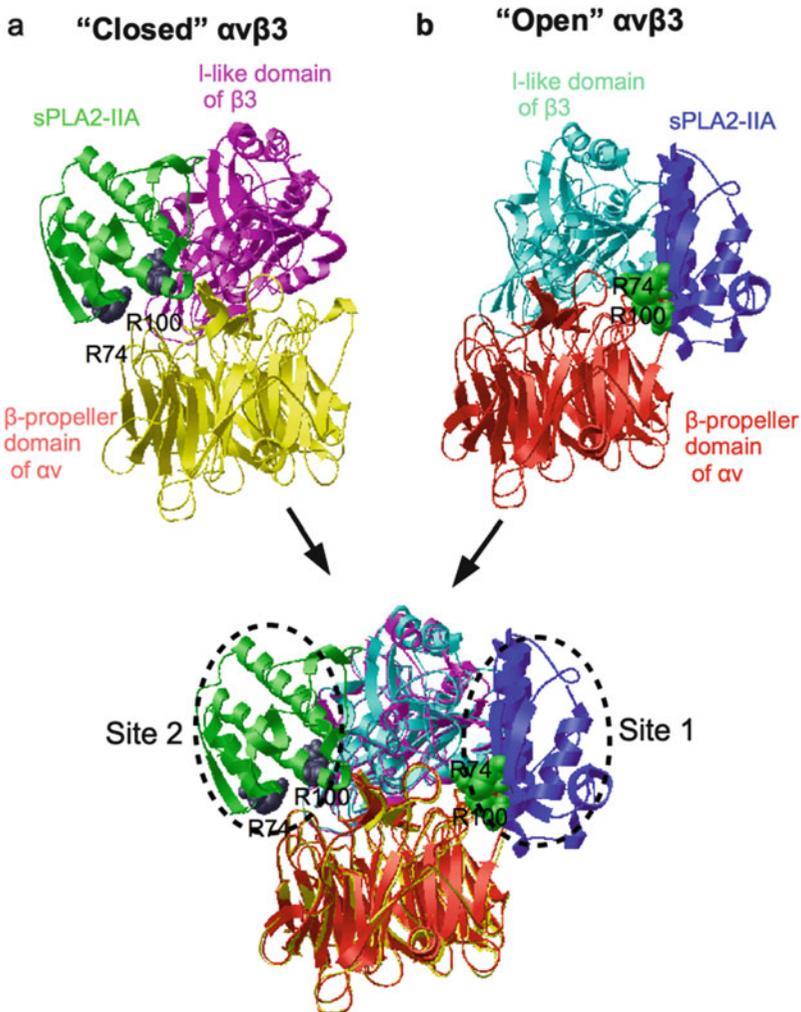
Fig. 3 Docking simulation predicts that sPLA2-IIA binds to a binding site that is distinct from the classical RGD-binding site in closed-head-piece αvβ3. We showed that sPLA2-IIA activates αvβ3 integrin in cell-free conditions (through direct integrin binding) (Fig. 2). If sPLA2-IIA and the fibrinogen fragment bind to the same binding site, sPLA2-IIA is expected to inhibit the binding of the fibrinogen fragment, but not enhance it. We suspected that the integrin activation by sPLA2-IIA may be mediated by the binding of sPLA2-IIA to another binding site. We also suspected that the second ligand-binding site (possible allosteric site) may be open in inactive integrins, but is closed in activated form. To predict the location of such a binding site, we performed

docking simulation using inactive αvβ3 as a target (1JV2. pdb). (**a**) a docking model of sPLA2-IIA-integrin αvβ3 (inactive) interaction. The headpiece of an inactive form of integrin αvβ3 (1JV2.pdb) was used as a target. (**b**) a docking model of sPLA2-IIA-integrin αvβ3 (active) inter-action (Saegusa et al. 2008). The headpiece of the ligand-bound form of integrin αvβ3 (1L5G.pdb) was used as a target. The model predicts that sPLA2-IIA (PDB code 1DCY) binds to the classical RGD-binding site of the integrin αvβ3 head- piece (*Site 1*). The model predicts the position of the second sPLA2-IIA-binding site (*Site 2*). (**c**) superposition of two models shows that the positions of two predicted sPLA2-IIA-binding sites are distinct

cells was maximum at 20 μg/ml sPLA2-IIA and then reduced as sPLA2-IIA concentration increases. This suggests that (1) sPLA2-IIA at low concentrations binds to site 2 of closed α4β1 (site 1 closed, site 2 open) and activates

α4β1 (site 1 open). (2) When site 2 is saturated with sPLA2-IIA, sPLA2-IIA competes with H120 for binding to site 1 (open) and reduces the binding of H120.

12 Biological Significance of sPLA2-IIA-Induced Integrin Activation

The present study establishes that sPLA2-IIA activates integrins through direct binding to site 2 (Fig. 4). sPLA2-IIA activated recombinant soluble αvβ3 in cell-free conditions, suggesting that inside-out signals or other molecules are not involved. This process does not include catalytic activity or the binding of sPLA2-IIA to M-type receptor since mutating the catalytic center (the H47Q mutation) or receptor-binding site of sPLA2-IIA (the G29S/D48K mutation) did not affect sPLA2-IIA-mediated integrin activation. sPLA2-IIA-induced integrin activation through binding to site 2 is a novel mechanism of integrin activation and signaling action by sPLA2-IIA. sPLA2-IIA may activate other integrins through direct binding.
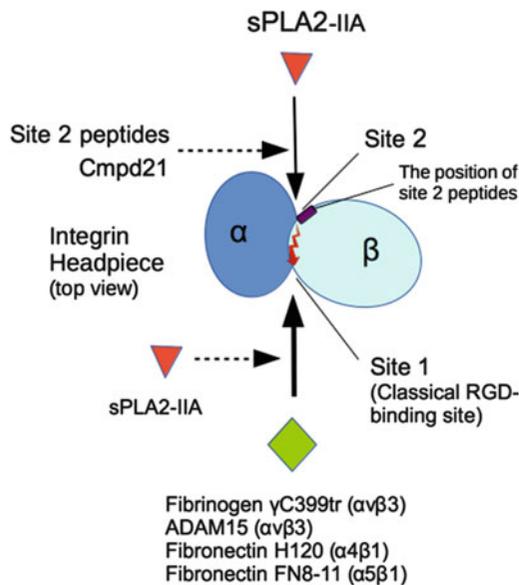

Fig. 4 A model of sPLA2-IIA-induced integrin activation through site 2. We propose a model, in which sPLA2-IIA binds to site 2 of inactive/closed-headpiece integrins and induces conformational changes and enhances ligand binding to site 1 (the classical RGD-binding site). This activation is blocked by a peptide that is derived from site 2 or a small compound (Cmpd21) that binds to the integrin-binding site of sPLA2-IIA

13 sPLA2-IIA Is Present at High Concentrations in Human Tears

sPLA2-IIA-mediated integrin activation happens in biological fluids (at least in tears). In our study, >5 μg/ml sPLA2-IIA is required to detect sPLA2-IIA-induced integrin activation. Notably, the concentration of sPLA2-IIA is exceptionally high in human tears (Nevalainen et al. 1994; Qu and Lehrer 1998; Birts et al. 2010). In normal subjects, the concentration of sPLA2-IIA in tears is 54.5 ± 33.9 μg/ml, one of the highest levels of sPLA2-IIA reported in any normal human secretions (Saari et al. 2001). Therefore integrin activation by sPLA2-IIA happens at least in tears. sPLA2-IIA appears to be secreted by both the lacrimal glands and the goblet cells of conjunctival epithelia (Nevalainen et al. 1994; Aho et al. 1996). Since sPLA2-IIA is bacteriocidal and kills Listera at much lower concentrations (<0.1 nM), it is possible that the primary functions of sPLA2-IIA at such high concentrations in tears might be integrin activation. It is likely that sPLA2-IIA in tears may play a role in enhancing immune response to bacterial pathogens through local integrin activation in tears or perhaps in other tissues. Serum levels of sPLA2-IIA are increased only up to 1 μg/ml during systemic inflammation (Jaross et al. 2002; Niessen et al. 2003). sPLA2-IIA may not effectively activate integrins at these concentrations. It is, however, possible that sPLA2-IIA may be highly concentrated in diseased tissues in chronic inflammation or on the cell surface through binding to proteoglycans.

14 sPLA2-IIA-Induced Integrin Activation Is Related to Its Signaling Action

The sPLA2-IIA-induced integrin activation is expected to enhance interaction between cells and extracellular matrix (e.g., fibrinogen and fibronectin) and thereby induce massive proliferative signals. Since integrins are involved

in growth factor signaling through crosstalk with growth factor receptors, sPLA2-IIA-induced integrin activation is also expected to enhance cellular responsiveness to growth factors. We have reported that integrins crosstalk with several growth factor receptors through direct binding to growth factors (e.g., fibroblast growth factor (FGF)-1 (Mori et al. 2008, 2013, Yamaji et al. 2010; Mori and Takada 2013), insulin-like growth factor (IGF)-1 (Saegusa et al. 2009; Fujita et al. 2012b, 2013a, b), neuregulin-1 (Ieguchi et al. 2010), and fractalkine (FKN) (Fujita et al. 2012a)). We propose that sPLA2-IIA-induced integrin activation indirectly affects intracellular signaling by these growth factors through enhancing integrin binding to growth factors.

15 sPLA2-IIA-Induced Integrin Activation Does Not Induce Global Conformational Changes

It is unclear if sPLA2-IIA-induced integrin activation requires global conformational changes in integrins. In current models of integrin activation, activation of β1 integrins induces a swing-out movement of the hybrid domain and exposes epitopes recognized by activation-dependent antibodies (anti-human β1 HUTS4 and HUTS21) (Luo et al. 2007). The HUTS4 and HUTS21 epitopes are located in the hybrid domain of β1 (Luque et al. 1996; Mould et al. 2003). sPLA2-IIA did not change reactivity of β1 integrins to HUTS4 and HUTS21 under the conditions in which sPLA2-IIA enhanced the binding of β1 integrins to ligands. It is thus possible that the binding of sPLA2-IIA to site 2 induces only local conformational changes within the headpiece of integrins.

Interestingly, the active (1L5G.pdb) and inactive (1JV2.pdb) conformations of αvβ3 are very similar (Xiong et al. 2001, 2002). Surprisingly, the docking simulation distinguished the two conformations and predicted the position of site 2 in the closed form. We showed that sPLA2-IIA and FKN-CD (Fujita et al. 2014) actually bind to

site 2 and the binding of sPLA2-IIA and FKN to site 2 is required for integrin activation. It is likely that integrins that are activated by sPLA2-IIA and FKN-CD through site 2 have conformations similar to the active αvβ3 that has no global conformational changes compared to the inactive form. The active and inactive conformations of αvβ3 may really reflect the fact that integrins can be activated without global conformational changes through allosteric mechanism.

16 Conclusion: sPLA2-IIA-Integrin Interaction Is a Novel Target for Drug Discovery

Catalytic activity of sPLA2-IIA has been a major target for drug discovery. Specific inhibitors of sPLA2-IIA catalytic activity S-5920/LY315920Na and S-3013/LY333013 failed to demonstrate a significant therapeutic effect in rheumatoid arthritis (Bradley et al. 2005) and asthma (Bowton et al. 2005). Our study suggests that sPLA2-IIA-integrin interaction may be a potential new target for chronic inflammatory diseases (Saegusa et al. 2008). Cmpd21 was screened for its ability to bind to the integrin-binding site of sPLA2-IIA (Ye et al. 2013). Indeed Cmpd21 suppressed the binding of sPLA2-IIA to integrins. Also, peptides from site 2 suppress sPLA2-IIA-induced integrin activation and have potential as therapeutics.

sPLA2-IIA is one of downstream effectors of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kB) (Dong et al. 2010; Antonio et al. 2002). Direct inhibition of NF-kB-mediated signal may be the most potent way to suppress inflammation in RA and sepsis shock. Indeed TNF-α blockers suppress NF-kB activation and work well in rheumatoid arthritis patients. TNF-α blockers, however, suppress general immune response of the patients and increase the risk of certain bacterial, mycobacterial, fungal, viral, and parasitic opportunistic pathogens and cancer (https://www.drugs.com/fda/tumor-necrosis-factor-alpha-tnf-alpha-blockers-label-change-boxed-warning-updated-risk-infection-13023.html;

Keane et al. 2001). We will be able to study if antagonists to sPLA2-IIA-integrin interaction such as Cmpd21 reduce chronic inflammation and if they suppress general immune response to pathogens in future studies.

Integrin activation through direct binding of sPLA2-IIA to an allosteric site (site 2) is potentially important for signaling functions of sPLA2-IIA. Since integrins are directly involved in signaling functions of several growth factors/cytokines (e.g., FGF, IGF, and fractalkine) (Chap. 14), we speculate integrin activation by sPLA2-IIA potentially enhances growth factor/cytokine signaling. Also, sPLA2-IIA-mediated integrin activation may enhance cell-extracellular matrix interaction. Allosteric activation of integrins through site 2 may be a novel target for drug discovery. Small molecular-weight compounds or peptides that block this process (e.g., site 2-derived peptides) may have potential as therapeutics. We will address this hypothesis in future studies.

References

Aho HJ, Saari KM, Kallajoki M, Nevalainen TJ (1996) Synthesis of group II phospholipase A2 and lysozyme in lacrimal glands. Invest Ophthalmol Vis Sci 37(9):1826–1832

Antonio V, Brouillet A, Janvier B, Monne C, Bereziat G, Andreani M, Raymondjean M (2002) Transcriptional regulation of the rat type IIA phospholipase A2 gene by cAMP and interleukin-1beta in vascular smooth muscle cells: interplay of the CCAAT/enhancer binding protein (C/EBP), nuclear factor-kappaB and Ets transcription factors. Biochem J 368(Pt 2):415–424

Birts CN, Barton CH, Wilton DC (2010) Catalytic and non-catalytic functions of human IIA phospholipase A2. Trends Biochem Sci 35(1):28–35. doi:10.1016/j.tibs.2009.08.003

Boilard E, Bourgoin SG, Bernatchez C, Poubelle PE, Surette ME (2003) Interaction of low molecular weight group IIA phospholipase A2 with apoptotic

human T cells: role of heparan sulfate proteoglycans. FASEB J 17(9):1068–1080. doi:10.1096/fj.02-0938com

Boilard E, Lai Y, Larabee K, Balestrieri B, Ghomashchi F, Fujioka D, Gobezie R, Coblyn JS, Weinblatt ME, Massarotti EM, Thornhill TS, Divangahi M, Remold H, Lambeau G, Gelb MH, Arm JP, Lee DM (2010) A novel anti-inflammatory role for secretory phospholipase A2 in immune complex-mediated arthritis. EMBO Mol Med 2 (5):172–187. doi:10.1002/emmm.201000072

Bowton DL, Dmitrienko AA, Israel E, Zeiher BG, Sides GD (2005) Impact of a soluble phospholipase A2 inhibitor on inhaled allergen challenge in subjects with asthma. J Asthma 42(1):65–71

Bradley JD, Dmitrienko AA, Kivitz AJ, Gluck OS, Weaver AL, Wiesenhutter C, Myers SL, Sides GD (2005) A randomized, double-blinded, placebo-controlled clinical trial of LY333013, a selective inhibitor of group II secretory phospholipase A2, in the treatment of rheumatoid arthritis. J Rheumatol 32 (3):417–423

Cierniewski CS, Byzova T, Papierak M, Haas TA, Niewiarowska J, Zhang L, Cieslak M, Plow EF (1999) Peptide ligands can bind to distinct sites in integrin alphaIIbbeta3 and elicit different functional responses. J Biol Chem 274(24):16923–16932

Cupillard L, Mulherkar R, Gomez N, Kadam S, Valentin E, Lazdunski M, Lambeau G (1999) Both group IB and group IIA secreted phospholipases A2 are natural ligands of the mouse 180-kDa M-type receptor. J Biol Chem 274(11):7043–7051

Dong Q, Patel M, Scott KF, Graham GG, Russell PJ, Sved P (2006) Oncogenic action of phospholipase A2 in prostate cancer. Cancer Lett 240(1):9–16. doi:10.1016/j.canlet.2005.08.012

Dong Z, Liu Y, Scott KF, Levin L, Gaitonde K, Bracken RB, Burke B, Zhai QJ, Wang J, Oleksowicz L, Lu S (2010) Secretory phospholipase A2-IIa is involved in prostate cancer progression and may potentially serve as a biomarker for prostate cancer. Carcinogenesis 31 (11):1948–1955

Dong Z, Meller J, Succop P, Wang J, Wikenheiser-Brokamp K, Starnes S, Lu S (2014) Secretory phospholipase A2-IIa upregulates HER/HER2-elicited signaling in lung cancer cells. Int J Oncol 45(3):978–984. doi:10.3892/ijo.2014.2486

Du XP, Plow EF, Frelinger AL 3rd, O'Toole TE, Loftus JC, Ginsberg MH (1991) Ligands "activate" integrin alpha IIb beta 3 (platelet GPIIb-IIIa). Cell 65(3):409–416

Fujita M, Takada YK, Takada Y (2012a) Integrins alphavbeta3 and alpha4beta1 act as coreceptors for fractalkine, and the integrin-binding defective mutant of fractalkine is an antagonist of CX3CR1. J Immunol 189(12):5809–5819. doi:10.4049/jimmunol.1200889

Fujita M, Ieguchi K, Davari P, Yamaji S, Taniguchi Y, Sekiguchi K, Takada YK, Takada Y (2012b) Crosstalk between integrin alpha6beta4 and insulin-like

growth factor-1 receptor (IGF1R) through direct alpha6beta4 binding to IGF1 and subsequent alpha6beta4-IGF1-IGF1R ternary complex formation in anchorage-independent conditions. J Biol Chem 287(15):12491–12500. doi:10.1074/jbc.M111.304170

Fujita M, Ieguchi K, Cedano-Prieto DM, Fong A, Wilkerson C, Chen JQ, Wu M, Lo SH, Cheung AT, Wilson MD, Cardiff RD, Borowsky AD, Takada YK, Takada Y (2013a) An integrin binding-defective mutant of insulin-like growth factor-1 (R36E/R37E IGF1) acts as a dominant-negative antagonist of the IGF1 receptor (IGF1R) and suppresses tumorigenesis but still binds to IGF1R. J Biol Chem 288 (27):19593–19603. doi:10.1074/jbc.M113.470872

Fujita M, Takada YK, Takada Y (2013b) Insulin-like Growth Factor (IGF) signaling requires alphavbeta3-IGF1-IGF type 1 receptor (IGF1R) ternary complex formation in anchorage independence, and the complex formation does not require IGF1R and Src activation. J Biol Chem 288(5):3059–3069. doi:10.1074/jbc.M112.412536

Fujita M, Takada YK, Takada Y (2014) The chemokine fractalkine can activate integrins without CX3CR1 through direct binding to a ligand-binding site distinct from the classical RGD-binding site. PLoS One 9(5): e96372. doi:10.1371/journal.pone.0096372

Fujita M, Zhu K, Fujita CK, Zhao M, Lam KS, Kurth MJ, Takada YK, Takada Y (2015) Proinflammatory secreted phospholipase A2 type IIA (sPLA-IIA) induces integrin activation through direct binding to a newly identified binding site (site 2) in integrins alphavbeta3, alpha4beta1, and alpha5beta1. J Biol Chem 290(1):259–271. doi:10.1074/jbc.M114.579946

Gelb MH, Cho W, Wilton DC (1999) Interfacial binding of secreted phospholipases A(2): more than electrostatics and a major role for tryptophan. Curr Opin Struct Biol 9(4):428–432

Gelb MH, Valentin E, Ghomashchi F, Lazdunski M, Lambeau G (2000) Cloning and recombinant expression of a structurally novel human secreted phospholipase A2. J Biol Chem 275(51):39823–39826

Hu DD, White CA, Panzer-Knodle S, Page JD, Nicholson N, Smith JW (1999) A new model of dual interacting ligand binding sites on integrin alphaIIbbeta3. J Biol Chem 274(8):4633–4639

Hynes RO (2002) Integrins: bidirectional, allosteric signaling machines. Cell 110(6):673–687

Ieguchi K, Fujita M, Ma Z, Davari P, Taniguchi Y, Sekiguchi K, Wang B, Takada YK, Takada Y (2010) Direct binding of the EGF-like domain of neuregulin-1 to integrins ({alpha}v{beta}3 and {alpha}6{beta}4) is involved in neuregulin-1/ErbB signaling. J Biol Chem 285(41):31388–31398. doi:M110.113878 [pii] 10.1074/jbc.M110.113878

Jaross W, Eckey R, Menschikowski M (2002) Biological effects of secretory phospholipase A(2) group IIA on lipoproteins and in atherogenesis. Eur J Clin Investig 32(6):383–393

Jiang J, Neubauer BL, Graff JR, Chedid M, Thomas JE, Roehm NW, Zhang S, Eckert GJ, Koch MO, Eble JN,

Cheng L (2002) Expression of group IIA secretory phospholipase A2 is elevated in prostatic intraepithelial neoplasia and adenocarcinoma. Am J Pathol 160(2):667–671. doi:10.1016/S0002-9440(10)64886-9

Keane J, Gershon S, Wise RP, Mirabile-Levens E, Kasznica J, Schwieterman WD, Siegel JN, Braun MM (2001) Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent. N Engl J Med 345(15):1098–1104

Lambeau G, Ancian P, Barhanin J, Lazdunski M (1994) Cloning and expression of a membrane receptor for secretory phospholipases A2. J Biol Chem 269(3):1575–1578

Legler DF, Wiedle G, Ross FP, Imhof BA (2001) Superactivation of integrin alphavbeta3 by low antagonist concentrations. J Cell Sci 114(Pt 8):1545–1553

Luo BH, Carman CV, Springer TA (2007) Structural basis of integrin regulation and signaling. Annu Rev Immunol 25:619–647

Luque A, Gomez M, Puzon W, Takada Y, Sanchez-Madrid F, Cabanas C (1996) Activated conformations of very late activation integrins detected by a group of antibodies (HUTS) specific for a novel regulatory region (355–425) of the common beta 1 chain. J Biol Chem 271(19):11067–11075

Mori S, Takada Y (2013) Crosstalk between Fibroblast Growth Factor (FGF) receptor and integrin through direct integrin binding to FGF and resulting Integrin-FGF-FGFR ternary complex formation. Med Sci 1(1):20–36. doi:10.3390/medsci1010020

Mori S, Wu CY, Yamaji S, Saegusa J, Shi B, Ma Z, Kuwabara Y, Lam KS, Isseroff RR, Takada YK, Takada Y (2008) Direct binding of integrin {alpha}v {beta}3 to FGF1 plays a role in FGF1 signaling. J Biol Chem 283(26):18066–18075. doi:M801213200 [pii] 10.1074/jbc.M801213200

Mori S, Tran V, Nishikawa K, Kaneda T, Hamada Y, Kawaguchi N, Fujita M, Takada YK, Matsuura N, Zhao M, Takada Y (2013) A dominant-negative FGF1 mutant (the R50E mutant) suppresses tumorigenesis and angiogenesis. PLoS One 8(2): e57927. doi:10.1371/journal.pone.0057927

Mould AP, Barton SJ, Askari JA, McEwan PA, Buckley PA, Craig SE, Humphries MJ (2003) Conformational changes in the integrin beta A domain provide a mechanism for signal transduction via hybrid domain movement. J Biol Chem 278(19):17028–17035

Murakami M, Kambe T, Shimbara S, Yamamoto S, Kuwata H, Kudo I (1999) Functional association of type IIA secretory phospholipase A(2) with the glycosylphosphatidylinositol-anchored heparan sulfate proteoglycan in the cyclooxygenase-2-mediated delayed prostanoid-biosynthetic pathway. J Biol Chem 274(42):29927–29936

Nevalainen TJ, Aho HJ, Peuravuori H (1994) Secretion of group 2 phospholipase A2 by lacrimal glands. Invest Ophthalmol Vis Sci 35(2):417–421

Nicolas JP, Lambeau G, Lazdunski M (1995) Identification of the binding domain for secretory

phospholipases A2 on their M-type 180-kDa membrane receptor. J Biol Chem 270(48):28869–28873

Niessen HW, Krijnen PA, Visser CA, Meijer CJ, Erik Hack C (2003) Type II secretory phospholipase A2 in cardiovascular disease: a mediator in atherosclerosis and ischemic damage to cardiomyocytes? Cardiovasc Res 60(1):68–77

Qu XD, Lehrer RI (1998) Secretory phospholipase A2 is the principal bactericide for staphylococci and other gram-positive bacteria in human tears. Infect Immun 66(6):2791–2797

Saari KM, Aho V, Paavilainen V, Nevalainen TJ (2001) Group II PLA(2) content of tears in normal subjects. Invest Ophthalmol Vis Sci 42(2):318–320

Saegusa J, Akakura N, Wu CY, Hoogland C, Ma Z, Lam KS, Liu FT, Takada YK, Takada Y (2008) Pro-inflammatory secretory phospholipase A2 type IIA binds to integrins alphavbeta3 and alpha4beta1 and induces proliferation of monocytic cells in an integrin-dependent manner. J Biol Chem 283 (38):26107–26115. doi:M804835200 [pii] 10.1074/jbc.M804835200

Saegusa J, Yamaji S, Ieguchi K, Wu CY, Lam KS, Liu FT, Takada YK, Takada Y (2009) The direct binding of insulin-like growth factor-1 (IGF-1) to integrin alphavbeta3 is involved in IGF-1 signaling. J Biol Chem 284(36):24106–24114. doi:M109.013201 [pii] 10.1074/jbc.M109.013201

Sartipy P, Johansen B, Gasvik K, Hurt-Camejo E (2000) Molecular basis for the association of group IIA phospholipase A(2) and decorin in human atherosclerotic lesions. Circ Res 86(6):707–714

Six DA, Dennis EA (2000) The expanding superfamily of phospholipase A(2) enzymes: classification and characterization. Biochim Biophys Acta 1488(1–2):1–19

Sved P, Scott KF, McLeod D, King NJ, Singh J, Tsatralis T, Nikolov B, Boulas J, Nallan L, Gelb MH, Sajinovic M, Graham GG, Russell PJ, Dong Q (2004) Oncogenic action of secreted phospholipase A2 in prostate cancer. Cancer Res 64(19):6934–6940. doi:10.1158/0008-5472.CAN-03-3018

Tada K, Murakami M, Kambe T, Kudo I (1998) Induction of cyclooxygenase-2 by secretory phospholipases A2 in nerve growth factor-stimulated rat serosal mast cells is facilitated by interaction with fibroblasts and mediated by a mechanism independent of their enzymatic functions. J Immunol 161(9):5008–5015

Takada Y, Ye X, Simon S (2007) The integrins. Genome Biol 8(5):215

Tatulian SA (2001) Toward understanding interfacial activation of secretory phospholipase A2 (PLA2): membrane surface properties and membrane-induced structural changes in the enzyme contribute synergistically to PLA2 activation. Biophys J 80(2):789–800

Triggiani M, Granata F, Balestrieri B, Petraroli A, Scalia G, Del Vecchio L, Marone G (2003) Secretory phospholipases A2 activate selective functions in human eosinophils. J Immunol 170(6):3279–3288

Vadas P, Stefanski E, Pruzanski W (1985) Characterization of extracellular phospholipase A2 in rheumatoid synovial fluid. Life Sci 36(6):579–587

Valentin E, Lambeau G (2000) Increasing molecular diversity of secreted phospholipases A(2) and their receptors and binding proteins. Biochim Biophys Acta 1488(1–2):59–70

Xiao T, Takagi J, Coller BS, Wang JH, Springer TA (2004) Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics. Nature 432(7013):59–67. doi:10.1038/nature02976

Xiong JP, Stehle T, Diefenbach B, Zhang R, Dunker R, Scott DL, Joachimiak A, Goodman SL, Arnaout MA (2001) Crystal structure of the extracellular segment of integrin alpha Vbeta3. Science 294(5541):339–345. doi:10.1126/science.1064535 1064535 [pii]

Xiong JP, Stehle T, Zhang R, Joachimiak A, Frech M, Goodman SL, Arnaout MA (2002) Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. Science 296 (5565):151–155

Xiong JP, Mahalingham B, Alonso JL, Borrelli LA, Rui X, Anand S, Hyman BT, Rysiok T, Muller-Pompalla D, Goodman SL, Arnaout MA (2009) Crystal structure of the complete integrin alphaVbeta3 ectodomain plus an alpha/beta transmembrane fragment. J Cell Biol 186(4):589–600. doi:10.1083/jcb.200905085

Yamaji S, Saegusa J, Ieguchi K, Fujita M, Takada YK, Takada Y (2010) A novel fibroblast growth factor-1 (FGF1) mutant that acts as an FGF antagonist. PLoS One 5(4):e10273. doi:10.1371/journal.pone.0010273

Ye L, Dickerson T, Kaur H, Takada YK, Fujita M, Liu R, Knapp JM, Lam KS, Schore NE, Kurth MJ, Takada Y (2013) Identification of inhibitors against interaction between pro-inflammatory sPLA2-IIA protein and integrin alphavbeta3. Bioorg Med Chem Lett 23 (1):340–345. doi:10.1016/j.bmcl.2012.10.080

Yokoyama K, Zhang XP, Medved L, Takada Y (1999) Specific binding of integrin alpha v beta 3 to the fibrinogen gamma and alpha E chain C-terminal domains. Biochemistry 38(18):5872–5877

Yokoyama K, Erickson HP, Ikeda Y, Takada Y (2000) Identification of amino acid sequences in fibrinogen gamma-chain and tenascin C C-terminal domains critical for binding to integrin alpha vbeta 3. J Biol Chem 275(22):16891–16898. doi:10.1074/jbc.M000610200, M000610200 [pii]

Zhu J, Zhu J, Springer TA (2013) Complete integrin headpiece opening in eight steps. J Cell Biol 201(7):1053–1068. doi:10.1083/jcb.201212037

The Simple and Unique Allosteric Machinery of *Thermus caldophilus* Lactate Dehydrogenase

Structure-Function Relationship in Bacterial Allosteric LDHs

Hayao Taguchi

Abstract

Many bacterial ʟ-lactate dehydrogenases (LDH) are allosteric enzymes, and usually activated by fructose 1,6-bisphosphate (FBP) and often also by substrate pyruvate. The active and inactive state structures demonstrate that *Thermus caldophilus*, *Lactobacillus casei*, and *Bifidobacterium longum* LDHs consistently undergo allosteric transition according to Monod-Wyman-Changeux model, where the active (R) and inactive (T) states of the enzymes coexist in an allosteric equilibrium (pre-existing equilibrium) independently of allosteric effectors. The three enzymes consistently take on open and closed conformations of the homotetramers for the T and R states, coupling the quaternary structural changes with the structural changes in binding sites for substrate and FBP though tertiary structural changes. Nevertheless, the three enzymes undergo markedly different structural changes from one another, indicating that there is a high variety in the allosteric machineries of bacterial LDHs. *L. casei* LDH undergoes the largest quaternary structural change in the three enzymes, and regulates its catalytic activity though a large linkage frame for allosteric motion. In contrast, *T. caldophilus* LDH exhibits the simplest allosteric motion in the three enzymes, involving a simple mobile structural core for the allosteric motion. TcLDH likely mediates its allosteric equilibrium mostly through electrostatic repulsion within the protein molecule, providing an insight for regulation machineries in bacterial allosteric LDHs.

H. Taguchi (✉)
Department of Applied Biological Science, Faculty of
Science and Technology, Tokyo University of Science,
2641 Yamazaki, Noda, Chiba 278-8510, Japan
e-mail: htaguchi@rs.noda.tus.ac.jp

Abbreviations

BlLDH	*Bifidobacterium longum* LDH
DMLDH	dogfish muscle LDH
FBP	fructose 1,6-bisphosphate
GsLDH	*Geobacillus stearothermophilus* LDH
LcLDH	*Lactobacillus casei* LDH
LDH	L-lactate dehydrogenase
LpLDH	*Lactobacillus pentosus* LDH
RMSD	root mean square deviation
TcLDH	*Thermus caldophilus* LDH
TmLDH	*Thermotoga maritima* LDH
TtLDH	*Thermus thermophilus* LDH
TtMDH	*Thermus thermophilus* MDH

L-Lactate dehydrogenase (LDH; EC 1.1.1.27) catalyzes oxidation-reduction between pyruvate and L-lactate with a coenzyme, NAD, and usually acts at the final step of anaerobic glycolysis to regenerate NAD^+ that is required for glycolysis (Holbrook et al. 1975). LDH forms a tetrameric structure comprising four identical subunits, which are related through three two-fold axes, the P, Q and R axes. Although LDH is one of the most studied enzymes for protein structure and structure-function relationship, it is also one of the highly divergent enzymes in primary structure and catalytic properties, depending on the organism or tissue. Figure 1 shows amino acid alignment for representative LDHs that are mentioned in this review. These LDHs show only about 40 % identity to one another, except for *Lactobacillus case* LDH (LcLDH) and *L. pentosus* LDH (LpLDH), which show 67 % amino acid identity though these two enzymes exhibit greatly different catalytic properties (Taguchi and Ohta 1992). Whereas the vertebrates have non-allosteric types of LDH isozymes, such as muscle-type and heart-type LDHs, which exhibit different catalytic properties (i.e. different substrate K_m)

from each other (Holbrook et al. 1975), many bacteria possess allosteric types of LDH, which are commonly activated by fructose 1,6-bisphosphate (FBP) (Garvie 1980). The FBP-mediated regulation for LDH activity is physiologically reasonable, since FBP is accumulated when NAD^+/NADH ratio is decreased within the bacterial cells.

Allostery is one of the important properties of proteins that plays a crucial role in the regulation of many biological processes, such as metabolism and signal transduction in the cell, but is also one of the complicated phenomena for experimental analysis. In 1960s, two representative models, the MWC model proposed by Monod et al. (1965) and the KNF model by Koshland et al. (1966), were proposed to explain the allosteric phenomenon. The former model is based on the pre-existing equilibrium, which is expanded from the lock and key theory (Fischer 1894), and the latter is based on the induced-fit theory (Koshland 1958). Nevertheless, the concepts and models for allostery have been greatly developed in the last few decades as the markedly improved experimental technologies, revealing a high variety in the regulatory strategies of allosteric proteins (Motlagh et al. 2014).

The active and inactive structures of LcLDH (Arai et al. 2010), *Bifidobacterium longum* LDH (BlLDH) (Iwata et al. 1994) and *Thermus caldophilus* LDH (TcLDH) (Ikehara et al. 2014) have been determined, and provide insights of the regulation machineries in bacterial allosteric LDHs. The allosteric transitions of the three enzymes can be simply explained by the pre-existing allosteric equilibrium between the two major states of protein structure proposed in MWC model (Monod et al. 1965). Hence, the active (R) and inactive (T) state structures of the enzymes coexist in an allosteric equilibrium
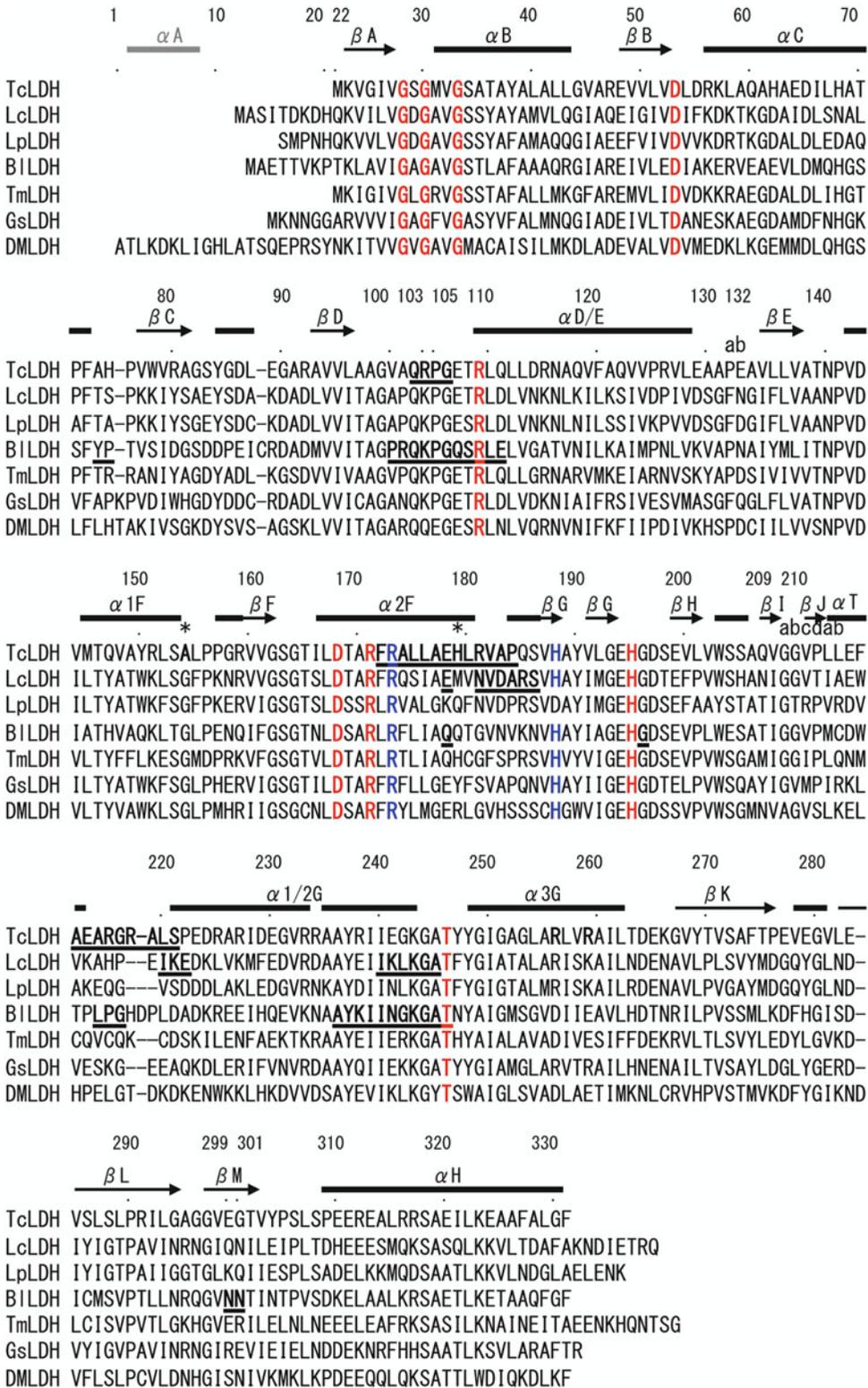
```
               1        10        20 22      30        40        50        60        70
                   α A                 β A          α B            β B          α C

TcLDH                                 MKVGIVGSGMVGSATAYALALLGVAREVVLVDLDRKLAQAHAEDILHAT
LcLDH               MASITDKDHQKVILVGDGAVGSSYAYAMVLQGIAQEIGIVDIFKDKTKGDAIDLSNAL
LpLDH                 SMPNHQKVVLVGDGAVGSSYAFAMAQQGIAEEFVIVDVVKDRTKGDALDLEDAQ
BlLDH             MAETTVKPTKLAVIGAGAVGSTLAFAAAQRGIAREIVLEDIAKERVEAEVLDMQHGS
TmLDH               MKIGIVGLGRVGSSTAFALLMKGFAREMVLIDVDKKRAEGDALDLIHGT
GsLDH             MKNNGGARVVVIGAGFVGASYVFALMNQGIADEIVLTDANESKAEGDAMDFNHGK
DMLDH     ATLKDKLIGHLATSQEPRSYNKITVVGVGAVGMACAISILMKDLADEVALVDVMEDKLKGEMMDLQHGS
```

```
              80        90      100 103 105 110        120       130 132       140
          β C          β D                α D/E                      β E
                                                                   ab

TcLDH   PFAH-PVWVRAGSYGDL-EGARAVVLAAGVAQRPGETRLQLLDRNAQVFAQVVPRVLEAAPEAVLLVATNPVD
LcLDH   PFTS-PKKIYSAEYSDA-KDADLVVITAGAPQKPGETRLDLVNKNLKILKSIVDPIVDSGFNGIFLVAANPVD
LpLDH   AFTA-PKKIYSGEYSDC-KDADLVVITAGAPQKPGESRLDLVNKNLNILSSIVKPVVDSGFDGIFLVAANPVD
BlLDH   SFYP-TVSIDGSDDPEICRDADMVVITAGPRQKPGQSRLELVGATVNILKAIMPNLVKVAPNAIYMLITNPVD
TmLDH   PFTR-RANIYAGDYADL-KGSDVVIVAAGVPQKPGETRLQLLGRNARVMKEIARNVSKYAPDSIVIVVTNPVD
GsLDH   VFAPKPVDIWHGDYDDC-RDADLVVICAGANQKPGETRLDLVDKNIAIFRSIVESVMASGFQGLFLVATNPVD
DMLDH   LFLHTAKIVSGKDYSVS-AGSKLVVITAGARQQEGESRLNLVQRNVNIFKFIIPDIVKHSPDCIILVVSNPVD
```

```
                 150       160       170       180       190       200      209 210
          α 1F          β F        α 2F            β G  β G      β H      β I β J αT
                  *              *                                          abcdab

TcLDH   VMTQVAYRLSALPPGRVVGSGTILDTARFRALLAEHLRVAPQSVHAYVLGEHGDSEVLVWSSAQVGGVPLLEF
LcLDH   ILTYATWKLSGFPKNRVVGSGTSLDTARFRQSIAEMVNVDARSVHAYIMGEHGDTEFPVWSHANIGGVTIAEW
LpLDH   ILTYATWKFSGFPKERVIGSGTSLDSSRLRVALGAFNVDPRSVDAYIMGEHGDSEFAAYSTATIGTRPVRDV
BlLDH   IATHVAQKLTGLPENQIFGSGTNLDSARLRFLIAQQTGVNVKNVHAYIAGEHGDSEVPLWESATIGGVPMCDW
TmLDH   VLTYFFLKESGMDPRKVFGSGTVLDTARLRTLIAQHCGFSPRSVHYYVIGEHGDSEVPVWSGAMIGGIPLQNM
GsLDH   ILTYATWKFSGLPHERVIGSGTILDTARFRFLLGEYFSVAPQNVHAYIIGEHGDTELPVWSQAYIGVMPIRKL
DMLDH   VLTYVAWKLSGLPMHRIIGSGCNLDSARFRYLMGERLGVHSSSCHGWVIGEHGDSSVPVWSGMNVAGVSLKEL
```

```
          220       230       240       250       260       270       280
                       α 1/2G              α 3G              β K

TcLDH   AEARGR-ALSPEDRARIDEGVRRAAYRIIEGKGATYYGIGAGLARLVRAILTDEKGVYTVSAFTPEVEGVLE-
LcLDH   VKAHP--EIKEDKLVKMFEDVRDAAYEIIKLKGATFYGIATALARISKAILNDENAVLPLSVYMDGQYGLND-
LpLDH   AKEQG---VSDDDLAKLEDGVRNKAYDIINLKGATFYGIGTALMRISKAILRDENAVLPVGAYMDGQYGLND-
BlLDH   TPLPGHDPLDADKREEIHQEVKNAAYKIINGKGATNYAIGMSGVDIIEAVLHDTNRILPVSSMLKDFHGISD-
TmLDH   CQVCQK--CDSKILENFAEKTKRAAYEIIERKGATHYAIALAVADIVESIFFDEKRVLTLSVYLEDYLGVKD-
GsLDH   VESKG--EEAQKDLERIFVNVRDAAYQIIEKKGATYYGIAMGLARVTRAILHNENAILTVSAYLDGLYGERD-
DMLDH   HPELGT-DKDKENWKKLHKDVVDSAYEVIKLKGYTSWAIGLSVADLAETIMKNLCRVHPVSTMVKDFYGIKND-
```

```
          290      299 301      310       320       330
          β L         β M            α H

TcLDH   VSLSLPRILGAGGVEGTVYPSLSPEEREALRRSAEILKEAAFALGF
LcLDH   IYIGTPAVINRNGIQNILEIPLTDHEEESMQKSASQLKKVLTDAFAKNDIETRQ
LpLDH   IYIGTPAIIGGTGLKQIIESPLSADELKKMQDSAATLKKVLNDGLAELENK
BlLDH   ICMSVPTLLNRQGVNNTINTPVSDKELAALKRSAETLKETAAQFGF
TmLDH   LCISVPVTLGKHGVERILELNLNEEELEAFRKSASILKNAINEITAEENKHQNTSG
GsLDH   VYIGVPAVINRNGIREVIEIELNDDEKNRFHHSAATLKSVLARAFTR
DMLDH   VFLSLPCVLDNHGISNIVKMKLKPDEEQQLQKSATTLWDIQKDLKF
```

Fig. 1 Structure-based sequence alignment of representative L-LDHs. The amino acid sequences of L-LDHs from seven organisms are aligned; *Thermus caldophilus* (TcLDH), *Lactobacillus casei* (LcLDH), *Lactobacillus*

(pre-existing equilibrium) independently of allosteric effectors, FBP or substrate, which possesses much higher affinity to the R state enzyme than the T state enzymes. Nevertheless, the three enzymes undergo markedly different structural changes in the allosteric transition, implying that bacterial allosteric LDHs have highly diverged their regulation machinery during evolution. This review mainly describes the comparison of the structural changes of these three enzymes in allosteric transition. TcLDH exhibits the smallest and simplest structural change in the three enzymes, showing its simple and unique allosteric machinery in the bacterial allosteric LDHs.

1 Catalytic Properties of Bacterial Allosteric Enzymes

Bacterial allosteric LDH exhibit highly various kinetic properties in the substrate K_m, k_{cat}, and FBP-dependency (Garvie 1980). The properties are also highly variable, depending on assay conditions such as pH and temperature, and minor structural change such as one point mutation. FBP usually induces a drastic reduction of substrate K_m ($S_{0.5}$), often together with a slight increase of V_{max} in the enzymes. Many of the enzymes show sigmoidal shapes of pyruvate saturation curves in the absence of FBP, indicating that the substrate exhibits homotropic activation effects on the enzyme reaction. Nevertheless, some bacterial allosteric LDHs show hyperbolic pyruvate saturations in the absence of FBP, which only improves the substrate K_m.

TcLDH shows a slightly sigmoidal shape of pyruvate saturation curve (Hill's coefficient 1.4), and gives 36 mM of substrate $S_{0.5}$ in the absence of FBP at 30 °C (Ikehara et al. 2014) (Fig. 2).

In the presence of FBP, this enzyme exhibits hyperbolic pyruvate saturation, and gives a 10^3-fold smaller substrate K_m. TcLDH significantly reduces the FBP-dependence at physiological temperature for the host organism, i.e. about 10-fold degreases the substrate $S_{0.5}$ at 80 °C in the presence of FBP, while the K_m in the presence of FBP does not markedly change. On the other hand, *Thermus thermophilus* LDH (TtLDH) differs from TcLDH in only two amino acid residues, Gly154 and Tyr179 for Ala154 and His179 of TcLDH, respectively. Nevertheless, TtLDH exhibits only a 10-fold reduced substrate K_m in the absence of FBP, while it shows a comparative substrate K_m to that of TcLDH in the presence of FBP (Colletier et al. 2012). LcLDH exhibits a clear sigmoidal shape of the pyruvate saturation curve (Hill's coefficient 2.0) and 20 mM of substrate $S_{0.5}$ (K_m) in the absence of FBP. In this case, FBP also changes the sigmoidal curve to the hyperbolic one, and 70-fold reduces substrate K_m at pH 5.0 (Arai et al. 2010, 2011). However, LcLDH shows great pH-dependence in the allosteric regulation, and exhibit no activity with conventional concentration of pyruvate at pH 7.0 unless FBP is present. LcLDH requires 10^5-fold higher concentration of FBP at pH 7.0, although some divalent cations such as Mn^{2+} reduce the FBP concentration required for the enzyme activation. In the presence of FBP, TcLDH and LcLDH consistently exhibit apparently higher V_{max} values, but also show great substrate inhibition at high concentration of pyruvate. Since LDH catalyzes the reaction essentially through a compulsory ordered bi-bi mechanism (Holbrook et al. 1975), pyruvate binds to the enzyme-NAD$^+$ complex before NAD$^+$ is dissociated from the enzyme, and thereby forms a non-productive ternary complex.

Fig. 1 (continued) pentosus (LpLDH), *Bifidobacterium longum* (BlLDH), *Thermotoga maritima* (TmLDH), *Geobacillus stearothermophilus* (GsLDH), and dogfish muscle (DMLDH). The residues are numbered according to the N-system proposed by Eventoff et al. (1977). Conserved amino acids important for the catalysis are colored *red*. Conserved Arg173 and His188 of the FBP binding sites are colored *blue*. The residues that exhibit

more than 2.0 Å deviation (RMSD at Cα) between the T and R state structures of TcLDH, LcLDH and BlLDH are *underlined*. The secondary structural elements of TcLDH, which are identical in the two state structures, are indicated by *bars* on the sequence. Ala154 and His179, that are replaced with Gly and Tyr in *T. thermophilus* LDH (TtLDH), respectively, are indicated by *asterisks*
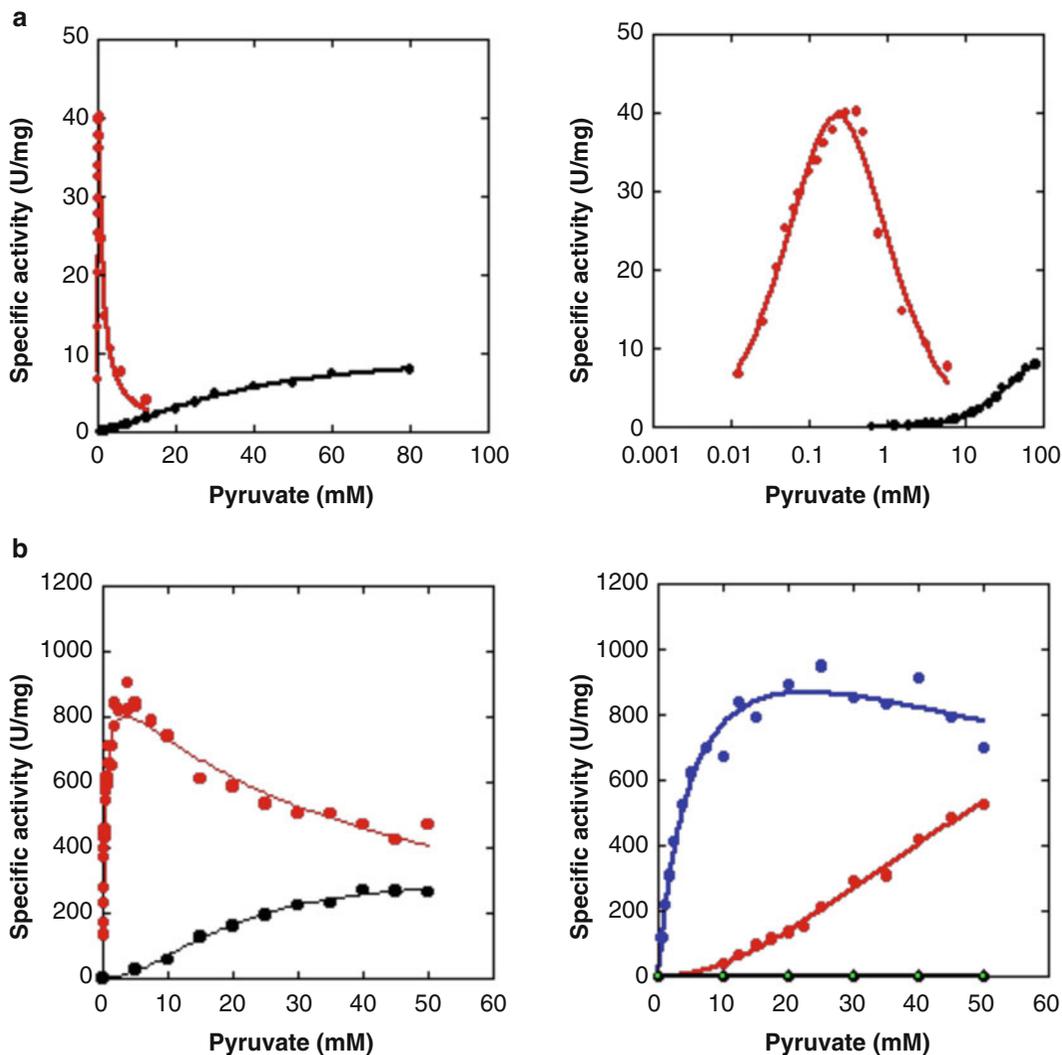
Fig. 2 Catalytic profiles of LcLDH and TcLDH at 30 °C. (**a**) pyruvate saturation curves for TcLDH without FBP (*black*) and with 0.1 mM FBP (*red*). The concentrations of pyruvate are plotted on linear (*left*) and logarithmic (*right*) scales. (**b**) pyruvate saturation curves for LcLDHs at pH 5.0 (*left*) and pH 7.0 (*right*) without FBP (*black*), with 10 mM MnCl$_2$ (*green*) or 5 mM FBP (*red*), and with 5 mM FBP and 10 mM MnCl$_2$ (*blue*)

Accordingly, the inhibition constant (K_i) for pyruvate is roughly correlated with pyruvate K_m, depending on the affinity to pyruvate (Eszes et al. 1996). The substrate inhibition explains apparently smaller V_{max} in the absence of FBP, where a high concentration of pyruvate simultaneously induces both the activation effects and inhibition effects on the enzyme reactions.

BlLDH exhibits a slight positive cooperativity (Hill's coefficient is 1.3) in pyruvate binding even in the presence of FBP, giving 0.2 mM of substrate $S_{0.5}$ (K_m). This enzyme shows a much higher positive cooperativity in pyruvate binding and an about 1000-fold larger substrate $S_{0.5}$ in the absence of FBP, although the exact kinetic parameters were not reported since the enzyme does not show the saturation profile at the conventional concentration of pyruvate (Fushinobu et al. 1996). In BlLDH, kinetic properties of subunit-hybrid enzymes were also reported (Fushinobu et al. 1996, 1998). The hybrid

BlLDHs comprise the subunits of the wild-type enzyme and mutant enzymes, which are modified in the FBP-binding site or the catalytic site, and show various catalytic profiles depending on the composition of the subunits. The hybridization analysis indicated that the structural changes of the active sites and substrate binding sites are cooperatively occur through the quaternary structural change in BlLDH, being consistent in the MWC model. On the other hand, the *Geobacillus stearothermophilus* enzyme (GsLDH), which is also well-studied for the structure and function in allosteric LDHs (Clarke et al. 1989), exhibits a hyperbolic saturation curve and gives a K_m value of 2 mM for

pyruvate without FBP, although the K_m value is 50-fold reduced in the presence of FBP (Clarke et al. 1986).

2   Quaternary Structural Changes

TcLDH, LcLDH and BlLDH consistently have expanded and contracted conformations of their quaternary structures for the T and R state structures, respectively (Fig. 3). The two structures of TcLDH are typical crystal structures for allosteric enzymes, since the R state structure is liganded with all the activating factors, FBP and oxamate (a pyruvate analogue), together
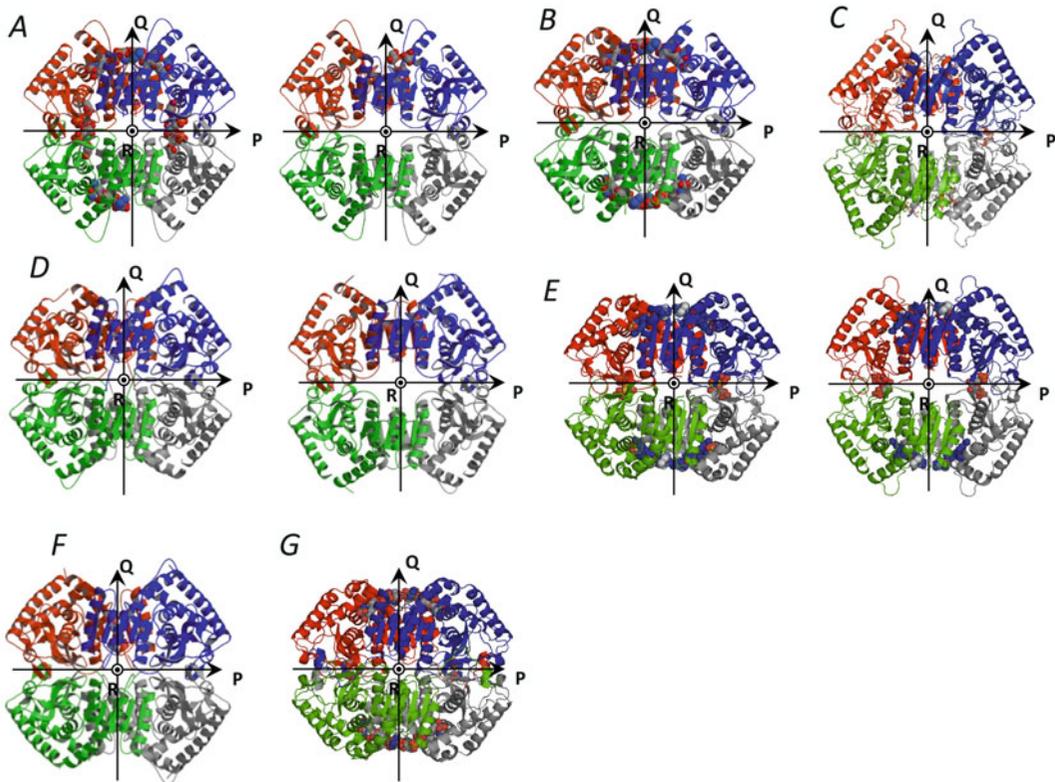
Fig. 3 Ribbon diagrams of the quaternary structures of the representative L-LDHs. (**a**) *left*, R state (liganded with NADH, oxamate and FBP) (PDB code 3VPH), *right*, T state (unliganded) (PDB 3VPG) structures of TcLDH. (**b**) R state (liganded with NADH and oxamate) TtLDH structure (PDB 2V7P). (**c**) liganded (with NADH, oxamate and FBP) TmLDH structure (PDB 1A5Z). (**d**) R (*left*) (PDB 2ZQZ) and T (*right*) state LcLDH structures, both of which are unliganded with

NADH, oxamate or FBP. (**e**) R state (liganded with FBP, NADH and oxamate) (*left*) and T state (liganded with FBP and NADH) (*right*) structures of BlLDH (PDB 1LTH). (**f**) non-allosteric LpLDH structure (unliganded) (PDB 1EZ4). (**g**) dogfish L-LDH structure (liganded with NADH and oxamate) (PDB 1LDM). All the structures are viewed along the molecular R-axis. The four subunits are colored *red*, *green*, *blue*, and *grey*. The P, Q, and R-axes are indicated by *arrows* and a *circle*

with NADH, whereas the T state structure is unliganded with any of the activators (Fig. 3a). The structures of the apo and holo (complex with NADH and oxamate) form enzymes were determined for *Thermus thermophilus* LDH (TtLDH) (Coquelle et al. 2007), which differs from TcLDH in only two amino acid residues, Gly154 and Tyr179 for Ala154 and His179 of TcLDH, respectively. The holo-form of TtLDH has essentially the same structure to the R state of TcLDH beside for the conformation of the active site loop region, in spite of lack of the bound FBP molecule (Fig. 3b). This is not surprising because oxamate is a potential activator, which exhibits homotropic activation effects on TcLDH instead of substrate pyruvate (Taguchi et al. 1985). The complex structure of *Thermotoga maritima* LDH (TmLDH), which has 45 % amino acid identity with TcLDH, was also determined (Auerbach et al. 1998). Although this structure contains FBP, NADH and oxamate, it resembles the T state structure of TcLDH rather than the R state structure, forming an expanded conformation of quaternary structure (Fig. 3c). This is also unsurprising, assuming that TmLDH exhibits the MWC-type allosteric transition. The structural comparison in the substrate and FBP binding sites also strongly suggests that this TmLDH structure is actually the T state structure of TmLDH, as described below.

It is particularly striking that the two state structures of LcLDH are consistently unliganded with any of the allosteric effectors (Fig. 3d), although the structures contain the bound SO_4 or NO_2 ions from the crystallization solvents, which mimic the phosphate groups of FBP (see Fig. 10c). This is the direct evidence indicating that LcLDH undergoes the allosteric transition that is explained by the pre-existence equilibrium theory proposed in MWC model. It is also notable that the two state structures of BlLDH, which are gained in a 1:1 ratio within the same crystal lattice, are consistently liganded with FBP (Iwata et al. 1994) (Fig. 3e). The T state BlLDH structure uncontaining FBP was also gained through a different crystalizing condition (Iwata and Ohta 1993). In this crystal, BlLDH binds NADH alone, and forms essentially the same (open) conformation as the FBP-bound T state enzyme.

The R state (closed) conformations of the three enzymes apparently resemble one another and also the structures of non-allosteric LDHs, such as *L. pentosus* LDH (LpLDH) (Fig. 3f) and the vertebrate LDHs (Fig. 3g), which constitutively exhibit high catalytic activities independent of FBP. It is noteworthy that the vertebrate enzymes have additional sequences at N-termini, which are extended by 20 amino acid residues from the first β-strand (βA) of the Rossmann-fold NAD-binding domain (Fig. 1). The extended sequences form inter-subunit interactions with the R-axis related subunit (Fig. 3g), and are called "the R-arms" (Abad-Zapetero et al. 1987), while these extended sequences are missing in the bacterial enzymes, particularly in TcLDH, which completely lacks the R-arm sequence. The R-arms apparently inhibit quaternary structural changes, and stabilize the contracted (R state) conformations for the vertebrate enzymes.

On the other hand, the three enzymes form markedly different conformations of the T state structures from one another, since they undergo different extents of structural change between the two states (Fig. 4 and Table 1). The P-axis-related dimers of the three enzymes consistently take on open and closed conformations between the two state structures by different magnitudes of motion (Fig. 4a). TcLDH and LcLDH exhibit the smallest and largest motions of the P-axis related subunits, respectively, and consequently the smallest and largest quaternary structural changes. Known LDHs have the largest inter-subunit contacts and many inter-subunit hydrogen bonds at the Q-axis subunit interface, where the αB and αC helices contact with the α2G-α3G area of the neighboring subunit. The three enzymes consistently form opened forms of Q-axis dimers in the T state structures, and retain most of these Q-axis contacts in the allosteric structural change, although BlLDH partially reconstructs the contacts and inter-subunit hydrogen bonds in the motion as described below. Unlike the vertebrate enzymes, which
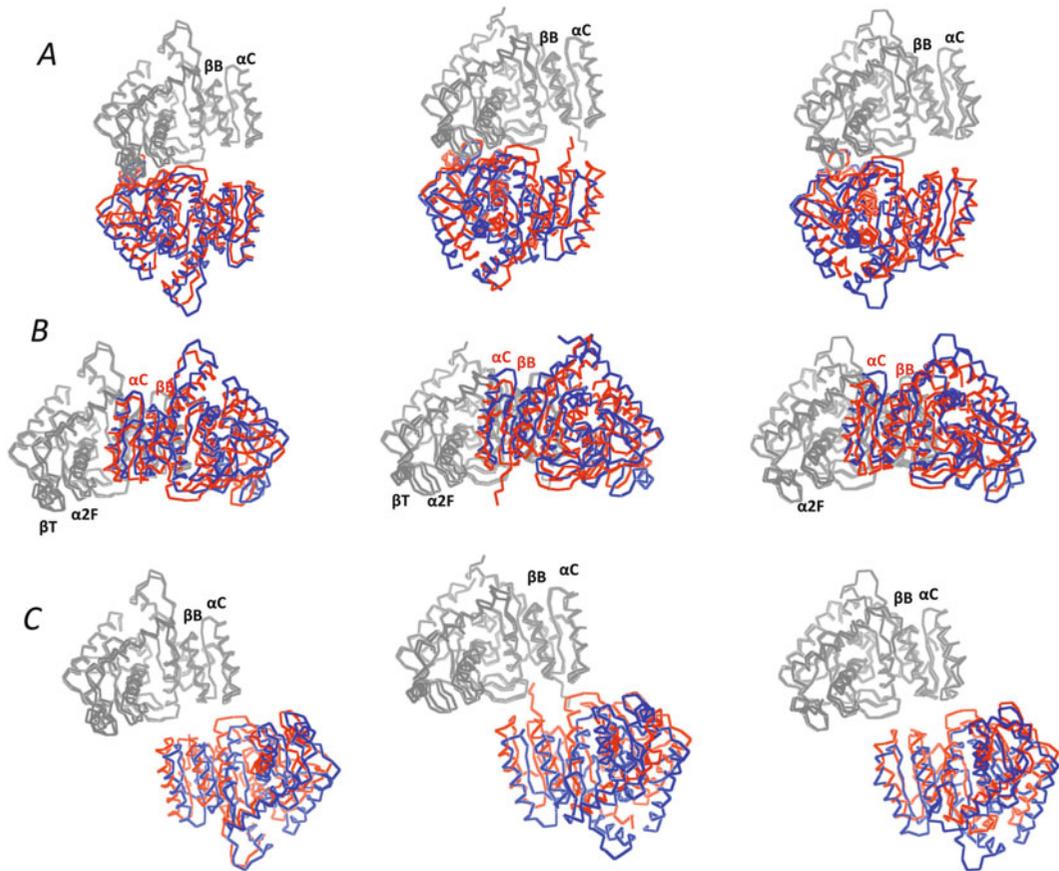
Fig. 4 Superimpositions of the R state (*red*) and T state (*blue*) structures of P-axis (**a**), Q-axis (**b**) and R-axis (**c**) related dimers for TcLDH (*left*), LcLDH (*center*), and BlLDH (*right*). The R and T state structures are superimposed by means of least squares deviation for a *grey* subunit of a dimer

Table 1 Quaternary structural changes in allosteric transition (average RSMD values at Cα atoms) (Å)[a]

		Monomer	P-dimer	Q-dimer	R-dimer	Tetramer
TcLDH(R)	TcLDH(T)	0.89	1.61	1.35	2.45	2.63
			(2.83)[c]	(3.10)[c]	(5.43)[c]	
TcLDH(R)	TtLDH(R)	0.79	0.60	0.69	0.52	0.76
without AS loop[b]		0.55	0.47	0.46	0.47	0.48
LcLDH(R)	LcLDH(T)	0.97	2.15	1.76	3.5	3.83
			(4.31)[c]	(4.07)[c]	(10.2)[c]	
BlLDH(R)	BlLDH(T)	1.26	1.87	1.9	2.75	2.98
			(3.52)[c]	(3.48)[c]	(6.5)[c]	

[a]The values were calculated using all Cα atoms available in whole structures, except for "without AS loop"
[b]The values were calculated without residues of the active site (AS) loop area (positions 98–109)
[c]The parentheses indicate the values for dimers that are fitted with only one subunit of dimer

consistently have the R-arm, the R state structures of the three enzymes form only small inter-subunit contacts and one or two inter-subunit hydrogen bonds at the R-axis subunit interfaces. The conformational changes to the T state structure further separate the two subunits from each other through the expanding motion of the quaternary structures, and completely remove the inter-subunit contacts and hydrogen bonds from the R-axis interface (Fig. 4 and Table 2).

Table 2 Numbers of inter-subunit hydrogen bonds and ion pairs

	TcLDH (R state)	TcLDH (T state)	LcLDH (R state)	LcLDH (T state)	BlLDH (R state)	BlLDH (T state)	LpLDH
H-bond							
P-axis	5	6	4	5	3	3	7
Q-axis	8	10	10	10	4	3	16
R-axis	2	0	1	0	1	0	3
Total	15	16	15	15	8	6	26
Ion pair							
P-axis	2	1	4	2	0	0	2
Q-axis	1	3	2	2	3	1	8
R-axis	0	0	1	0	0	0	1
Total	3	4	7	4	3	1	11

3 Tertiary Structural Changes

Known bacterial allosteric LDHs have virtually an identical secondary structural element composition to the vertebrate enzymes, although they consistently lack the αA helix located in the R-arm of the vertebrate enzymes (Fig. 1). Nevertheless, the region between βJ and α1/2G (positions 210a–220) is an exception. This region is called the flexible surface area, where both the primary and three-dimensional structures are poorly conserved in known LDHs. In the LDHs shown in Fig. 1, TcLDH, LcLDH, LpLDH and TmLDH have a small helix (the αT helix) at positions 210a-213 in the flexible surface area, while this helix is partially or completely unwound in BlLDH, GsLDH (Wigley et al. 1992) and dogfish muscle LDH (DMLDH, a representative vertebrate LDH) (Abad-Zapetero et al. 1987). TcLDH and LcLDH show no significant change in the secondary structural elements between the two state structures, and the α1/2G helix is constitutively kinked at position 234. On the other hand, BlLDH uniquely possesses the continuous α1/2G helix in its T state structure, whereas it has the kinked α1/2G helix in the R state structure as usual LDHs.

Like the vertebrate non-allosteric LDHs (Grau et al. 1981), allosteric LDHs exhibit great structural changes in the catalytic process, and the βD-αD/E loop (active site loop) area particularly takes on closed and open motion through the binding and releasing of ligands, respectively.

The R and T state structures of BlLDH have the closed and open forms of the active site loop, respectively, and therefore shows the greatest apparent structural change in the active site loop area between the two states contain the catalytic structural changes (Figs. 5 and 6). On the other hand, the two state structures of TcLDH and LcLDH consistently possess the open active site loops, though the loop of the R state TcLDH is slightly closed as compared with the one of the T state. The open form of the active site loop is quite unusual for the complex structures of LDH with a substrate analogue (usually oxamate). In the crystal structure of the R state TcLDH, the guanidium of Arg103 on the active site loop forms a salt-bridge with the Asp114 carboxyl group of the enzyme molecule belonging to the neighboring asymmetric unit. This salt-bridge interaction appears to stabilize the open conformation of the active site loop in the R state TcLDH. On the other hand, the holo (R state) structure of TtLDH has the closed form of the active site loop as usual. This conformational difference of the active site loop is mostly responsible for the slight structural difference between the R state TcLDH and TtLDH (Table 1). The structures of these enzymes thus clearly indicate that the allosteric structural changes occur independently of the active site loop motion, revealing net structural changes in the allosteric motions. In the allosteric motion, TcLDH, LcLDH and BlLDH consistently undergo obviously less and smaller structural changes in the NAD-binding domain (position

Fig. 5 Deviation of Cα atoms (RSMD values) between the R and T state structures. (**a**) TcLDH. (**b**) LcLDH (*blue*) and BlLDH (*green*). MR1, MR2, MR3 indicate mobile regions that display marked deviation (more than 2.0 Å) in the two state structures for any of TcLDH, LcLDH and BlLDH except for the active site loop area
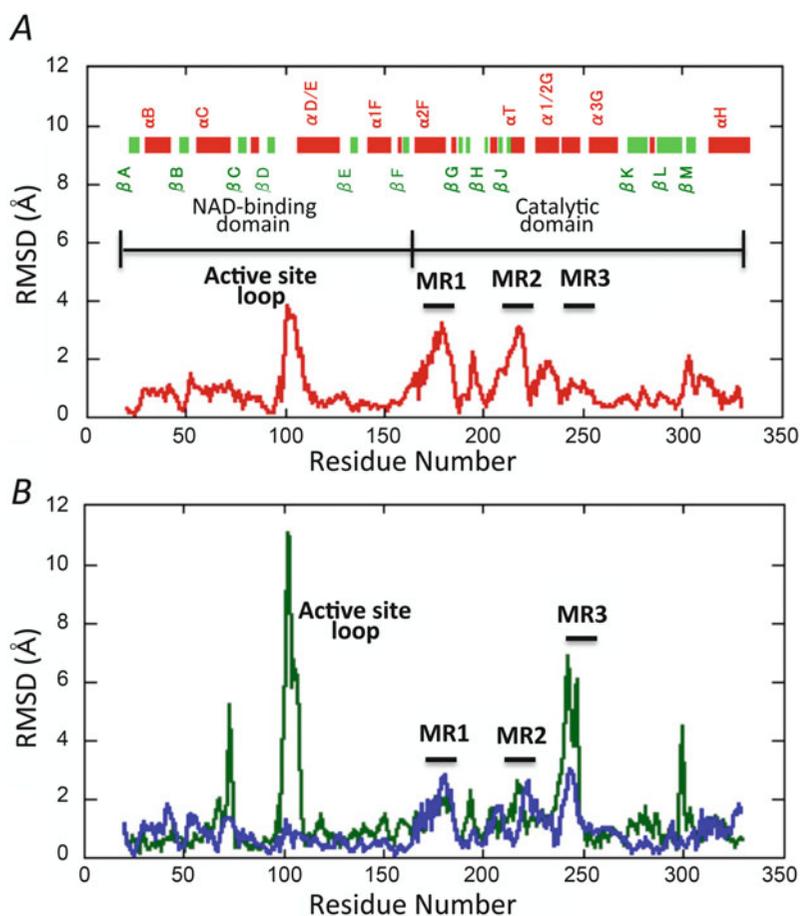
22–164) than the catalytic domain (165–330), beside for the active site loop (Fig. 5). In particular, the βA to βF strands constitute the rigid core β-sheet of the Rossman-fold domain, and move only slightly (Fig. 4). In this domain, nevertheless, TcLDH and LcLDH slightly (about 1 Å) move the αB and αC helices, which form intersubunit contacts with the α2-α3G area (mobile region 3 (MR3)) of the Q-axis related subunit in known LDHs. On the other hand, BlLDH does not significantly moves any of these two helices unlike the other two enzymes.

The three enzymes consistently exhibit large structural changes in the catalytic domain (166–331), but their profiles of motion are markedly different from one another (Figs. 5 and 6). TcLDH exhibits great structural changes in two regions (mobile regions; MRs), MR1 and MR2, which locate at the C-terminal area of the α2F

helix and the flexible surface area, respectively. The corresponding regions also markedly move in LcLDH and BlLDH, though less greatly than those of TcLDH. The two enzymes, particularly BlLDH, exhibit great structural changes in the α2G-α3G area (MR3), while TcLDH moves the MR3 only slightly. The three enzymes also markedly move the α1/2G helix area in different manners (Fig. 5). Superimpositioning of the T and R state structures clearly shows the different allosteric motions of the three LDHs (Fig. 6). The C-terminal area of the α2F helix (MR1) and the flexible surface area (MR2) form intra-subunit contact between them, and therefore move together. In particular, TcLDH greatly shifts these two regions in the same extent and same direction. In BlLDH, these two regions show smaller motions; instead the α1/2G-α3G loop area (MR3) greatly moves like the active site
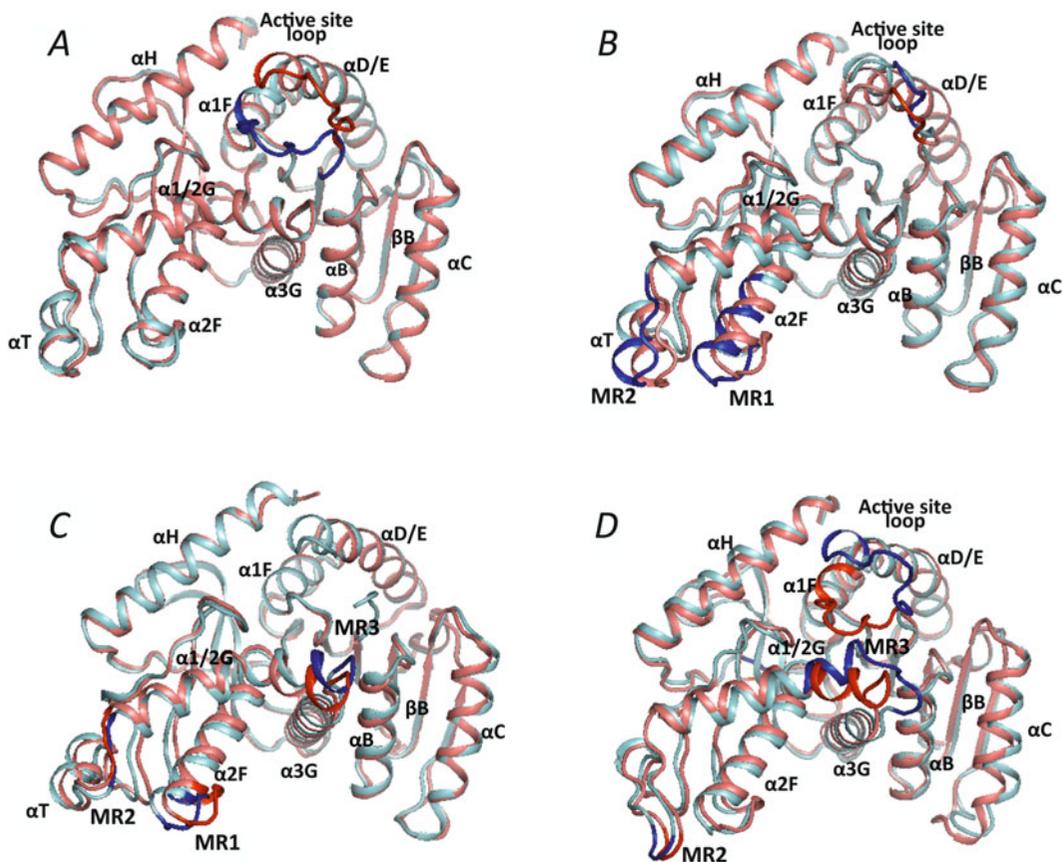
Fig. 6 Superimpositions showing tertiary structural changes. (**a**) the R state of TcLDH (*salmon red*), and R state TtLDH (*cyan*). (**b–d**) the R state (*salmon red*) and T state (*cyan*) structures of TcLDH (**b**), LcLDH (**c**) and BlLDH (**d**). The MR1 to MR3 regions are indicated by red and blue for the R and T states, respectively

loop. LcLDH shows an apparently intermediate type of structural change between those of TcLDH and BlLDH, both the MR1-MR2 and α1/2G-α3G areas moving in the equivalent magnitudes.

4   Linkage of Tertiary and Quaternary Structural Changes

MR1 and MR2 are located at the P-axis and Q-axis inter-subunit contact area, respectively, and thereby the motions of this area are closely coupled with the quaternary structural changes in the allosteric LDHs.

TcLDH, LcLDH and BlLDH form a main inter-subunit contact at the P-axis between the α2F-βG loop and the α2F-βG turn of the P-axis related subunit as in the cases of other LDHs. Since the α2F-βG loop is located at MR1 (Figs. 1 and 5), the MR1 motion is coupled with the quaternary structural change in the three enzymes (Fig. 7). In TcLDH and BlLDH, the MR1 moves together with the whole P-axis related subunit, and does not directly link to motion of any specific region of the neighboring subunit; i.e. the α3G helix of the counterpart subunit is unmoved in the tertiary structure of subunit. On the other hand, the MR1 motion does not only link to the quaternary structural change, but also links to the tertiary structural
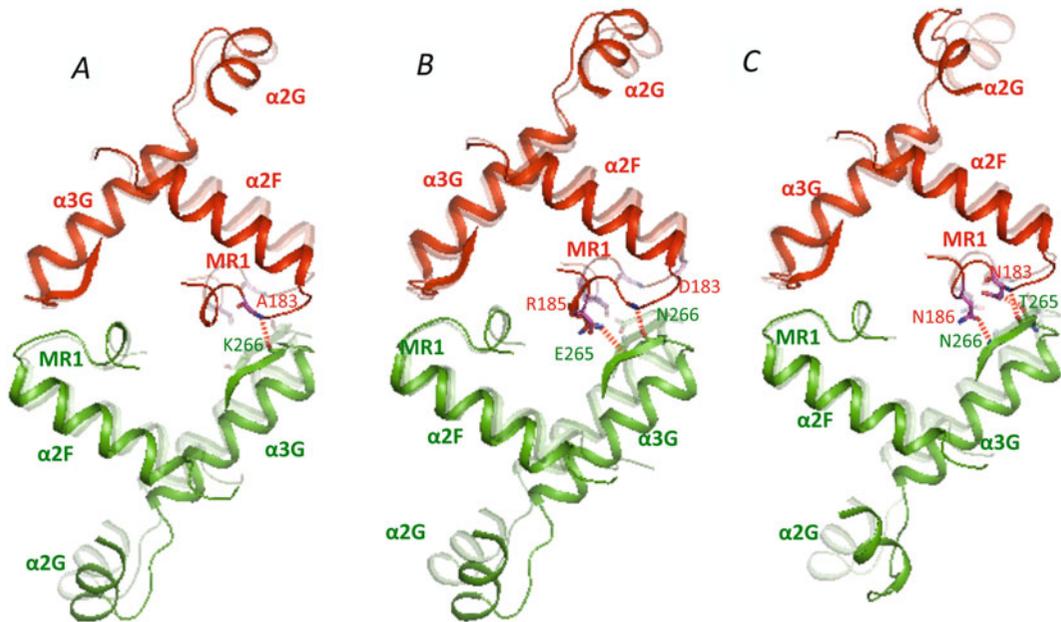
Fig. 7 Superimpositioning the R state (semi transparent) and T state structures of MR1 regions at P-axis subunit interfaces for TcLDH (**a**), LcLDH (**b**), and BlLDH (**c**). The *red* and *green* subunits are related at the P-axis subunit interfaces, and the two state structures are superimposed by means of least squares deviation for only one subunit (*red subunit*) of a dimer

change of the neighboring subunit in LcLDH, where the α3G helix slightly but significantly shifts along its along its helical axis (Fig. 7 center). In LcLDH, therefore, the MR1 motion forms inter-subunit linkage of the allosteric motion with the α3G(P) helix, which links the motion of the α2G-α3G(P) loop, i.e. the MR3(P).

MR3 is located at the C-terminal area of the α2G helix to the α2G-α3G loop, which forms inter-subunit contact with the αC helix of the Q-axis related subunit (Fig. 8). TcLDH and LcLDH form constitutive inter-subunit hydrogen bonds, which contain a salt-bridge between conserved Asp35 and Lys243, between the α2G-α3G loop and the αC(Q) helix in both the R and T state structures (Fig. 8a and b). On the other hand, BlLDH forms no inter-subunit hydrogen bond between the α2G-α3G loop and the αC (Q) helix in the T state structure, and only forms the inter-subunit in the R state structure (Fig. 8a and c). In the T state BlLDH structure, the α1/2G is a continuous straight helix, and Lys 243 on the α2G- α3G loop turns the orientation opposite to Asp65(Q) (Fig. 8c).

The α2G-α3G area (MR3) of TcLDH moves only slightly in the tertiary structure, because the flexible motions of the αB(Q) and αC(Q) helices sufficiently cushions the quaternary structural change for the inter-subunit contacts (Fig. 8d). The two helices move along their helical axes in the tertiary structure of their own subunit (the blue subunit in the figure), together with the whole neighboring subunit (the red subunit), according to the quaternary structural change. On the other hand, LcLDH more greatly changes the quaternary structure than TcLDH, and therefore the motions of the αB(Q) and αC(Q) alone insufficiently cushion the quaternary structural change. The α2G-α3G loop also moves together with the αC(Q) (Fig. 8e), keeping the inter-subunit hydrogen bonds, which are strengthen by Asp62(Q) on the αC(Q) helix (Fig. 8b) as compared to those of TcLDH (Fig. 8a). Concomitantly, the α3G helix slightly shift along the helical axis by a support of the flexible motion of the αB(Q) helix, forming inter-subunit linkage of motion between the P-axis related subunit (Fig. 7c). In the case of BlLDH, in contrast, the
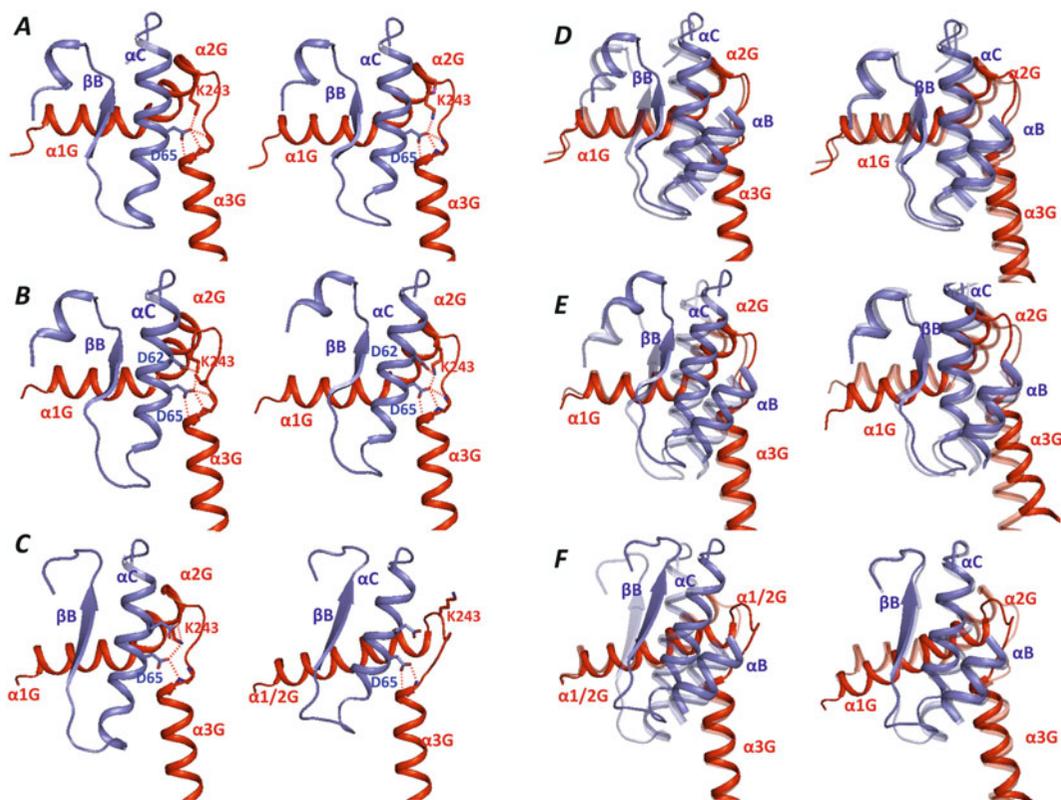
Fig. 8 The R and T state structures of MR3 regions at Q-axis subunit interfaces. (**a–c**) the R state (*left*), the T state (*center*) structures and their superimpositionings (*right*) for TcLDH (**a**) LcLDH (**b**) and BlLDH (**c**). The red and blue subunits are related at the Q-axis subunit interfaces. The *red dotted lines* indicate hydrogen bonds between the α2G-α3G area of *red* subunit and the αC helix of *blue* subunit. (**d–f**) superimpositioning of the T and R (*semitransparent*) state structures of TcLDH (**d**) LcLDH (**e**) and BlLDH (**f**). The two state structures are superimposed by means of least squares deviation for *red* subunit (*left*) and *blue* subunit (*right*)

αB(Q) and αC(Q) do not play the role of cushion for the quaternary structural change at all, instead the MR3 cushions the structural change by its flexible motion (Fig. 8f). These different structural changes in the Q-axis contacts directly lead to the different structural changes in the active sites of the three enzymes.

5   Structural Change in the Active Site

Substrate pyruvate directly interacts with the conserved His195, Arg109, Arg171 and Thr246 in the active site of LDH (Holbrook et al. 1975). The imidazole of His195 plays the acid/base catalyst, and provides the carbonyl oxygen of pyruvate with a proton, and the Arg171 guanidinium forms bidentate ionic hydrogen bonds with the carboxyl group of pyruvate to bind and to correctly orientate the substrate in the binding site (Hart et al. 1987). Thr246 also forms hydrogen bond with the carboxyl group of pyruvate to support the substrate binding and catalytic reaction of the enzyme (Bur et al. 1989; Sakowicz et al. 1993). On the other hand, Arg109 is located on the active site loop, and polarize the pyruvate carbonyl group to enhance the hydrogen transfer step in the catalytic reaction (Clarke et al. 1986). In addition, the conserved Asp168 also supports the catalytic function of LDH, forming hydrogen bonds with His195 (Clarke et al. 1988). The R state TcLDH structure is unique in the ternary complex
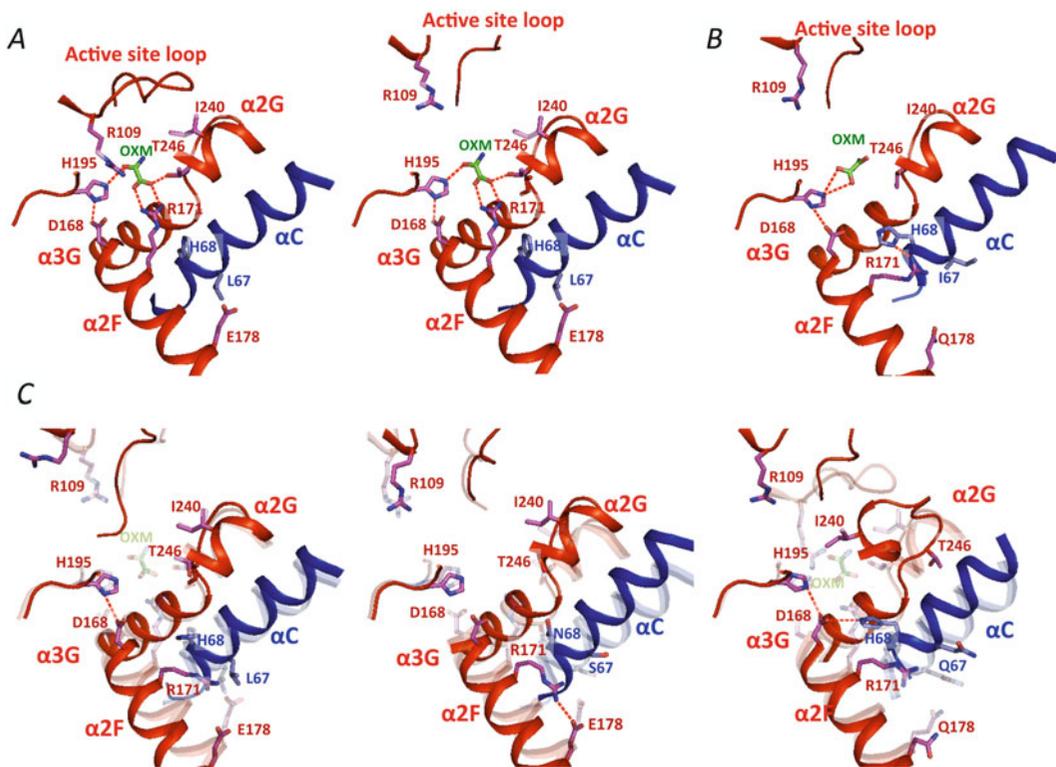
Fig. 9 Structural changes in the active site area. (**a**) the structures of the R state TtLDH (*left*) and the R state TcLDH (*right*). (**b**) the structure of TmLDH. (**c**) the superimpositioning of the T and R (*semitransparent*) state structures of TcLDH (*left*), LcLDH (*center*) and BlLDH (*right*). Oxm indicates the oxamate molecule bound in the enzymes, and *red dotted lines* indicate hydrogen bonds. The *red* and *blue* subunits are related at the Q-axis subunit interfaces, and the two state structures are superimposed by means of least squares deviation for Q-axis related dimers

(liganded with NADH and oxamate) structure that possesses the open form of the active site loop. In this structure, nevertheless, TcLDH forms the correct interactions with the bound oxamate molecule through all the relevant amino acid residues beside for Arg109, which locates at the open active site loop (Fig. 9a). This structure is essentially the same to the active site structure of holo TtLDH except for the conformation of the active site loop, providing a new structural snapshot of the catalytic process of LDHs. In the case of the TmLDH structure, in contrast, the oxamate molecule is bound to the enzyme in different orientation, and its carboxyl group form hydrogen bond interactions with His195, instead of Arg171, which orientates the guanidinium side chain outside the active site

(Fig. 9b), as in the cases of the T state structures of TcLDH, LcLDH and BlLDH (Fig. 9c). These evidences also indicate that TmLDH forms the T state structure in this crystal structure. The TmLDH crystal is prepared in the presence of 150 mM NADH, 250 mM oxamate, and 150 mM FBP. The unusual oxamate binding likely arose from such crystallization conditions that contain high concentrations of the ligands.

The three enzymes, TcLDH, LcLDH and BlLDH commonly orientate the side chain of Arg171 outside the active site in their T state structures, and switch the Arg171 orientation through a change in the contact between the α2F helix and the αC helix on the Q-axis related subunit (αC(Q)) (Fig. 9c), in which His68 (Q), or Asn68(Q) for LcLDH, blocks the alternative

orientation of Arg171 through steric hindrance. However, the three enzymes obviously differ from one another in the motions of the two helices. In TcLDH, the α2F helix greatly rotates around its N-terminal area, but the αC(Q) helix only slightly rotates around its C-terminal area (Fig. 9b, *left*). The large motion of the α2F helix and small motion of the αC(Q) helix arise from large MR1 motion (Fig. 7a) and small MR3 motion (Fig. 8d), respectively. The αC(Q) helix is tightly bound to the α2G-α3G loop (MR3) through constitutive inter-subunit hydrogen bonds, and moves as a part of the active site (red subunit in Figs. 7, 8, and 9) in the allosteric transition. Consequently, TcLDH only changes the Arg171 orientation in the active site, mostly through the α2F helix rotation around the N-terminal area of the helix, and does not markedly change the positions or orientations of His195, Thr246, or Asp168. On the other hand, LcLDH greatly rotates both the α2F and αC (Q) helices (Fig. 9b, *center*), accordingly to the large motions of MR1 (Fig. 7b) and MR3 (Fig. 8b), and concomitantly changes the position of Thr246 on the α2G-α3G loop (MR3) beside the change of Arg171 orientation. In contrast, BlLDH rotates the αC(Q) helix much more greatly than the α2F helix (Fig. 9b, *right*). This motion arises from the greater structural change of MR3 (Fig. 8f) than MR1 (Fig. 7c). Since the αC(Q) helix of BlLDH moves only together with the Q-axis related subunit (blue subunit in figure) without functioning as cushion for the quaternary structural change, it deeply intercalates into the active site in the T structure, allowing its His68(Q) to form hydrogen bonds with Asp168. In the α2G-α3G loop of the T state BlLDH, in addition, Thr246 completely point away from the active site, and Ile240 to occupy the position in which the substrate molecule would be accommodated in the R state. Hence, BlLDH blocks the active site from substrate pyruvate in the T state structure more completely than the other two enzymes, whereas the T state TcLDH has an apparently poorer guard for the active site than the other two enzymes.

6 Structural Change in the FBP Binding Site

Known allosteric LDHs consistently bind two FBP molecules per tetramer at the P-axis subunit interface in essentially the same manner (Iwata et al. 1994; Ikehara et al. 2014; Wigley et al. 1992). Each the two FBP molecules, which has a pseudo symmetric structure, is bound in each the two P-axis dimer units, where Arg173 and His188 of two juxtaposed subunits form salt-bridges with the two phosphate groups of FBP (Fig. 10). The R state TcLDH structure provides the FBP binding manner of the enzyme at the highest resolution. In the FBP binding site, the two His188 form multiple hydrogen bonds with both the two phosphate groups of FBP, while each of the two Arg173 forms monodentate or bidentate hydrogen bonds with each the two phosphates (Fig. 10a). Although the FBP binding site of TcLDH contains two glycerol molecules, these molecules appear to be only artifacts due to the crystallization solution. In the case of TcLDH, the two imidazolium rings of His188 are positioned in parallel, and orientated to the bound FBP molecule in the R state structure, but greatly point away from the binding site in the T state structure (Fig. 10b). It is important that the R state (holo) TtLDH has essentially the same His188 conformation as the R state TcLDH, although it contains no FBP molecule in the binding site. In the R state structure, hence, the His188 orientation suitable for binding FBP is not induced through interactions with the FBP molecule, but only pre-exists independently of the FBP molecule, being consistent with the MWC pre-existing theory. In the case of TcLDH or TtLDH, His188 plays the key role in the switch of the affinity to FBP between the two states through changing the orientation. In addition, the area of Gln185, which locates at or near the MR1 of the enzyme, markedly narrows the space of the FBP-binding site in the T state structure, and therefore likely prevents the FBP binding by steric hindrance.

The R and T state structures of LcLDH are consistently unliganded with FBP, instead bind
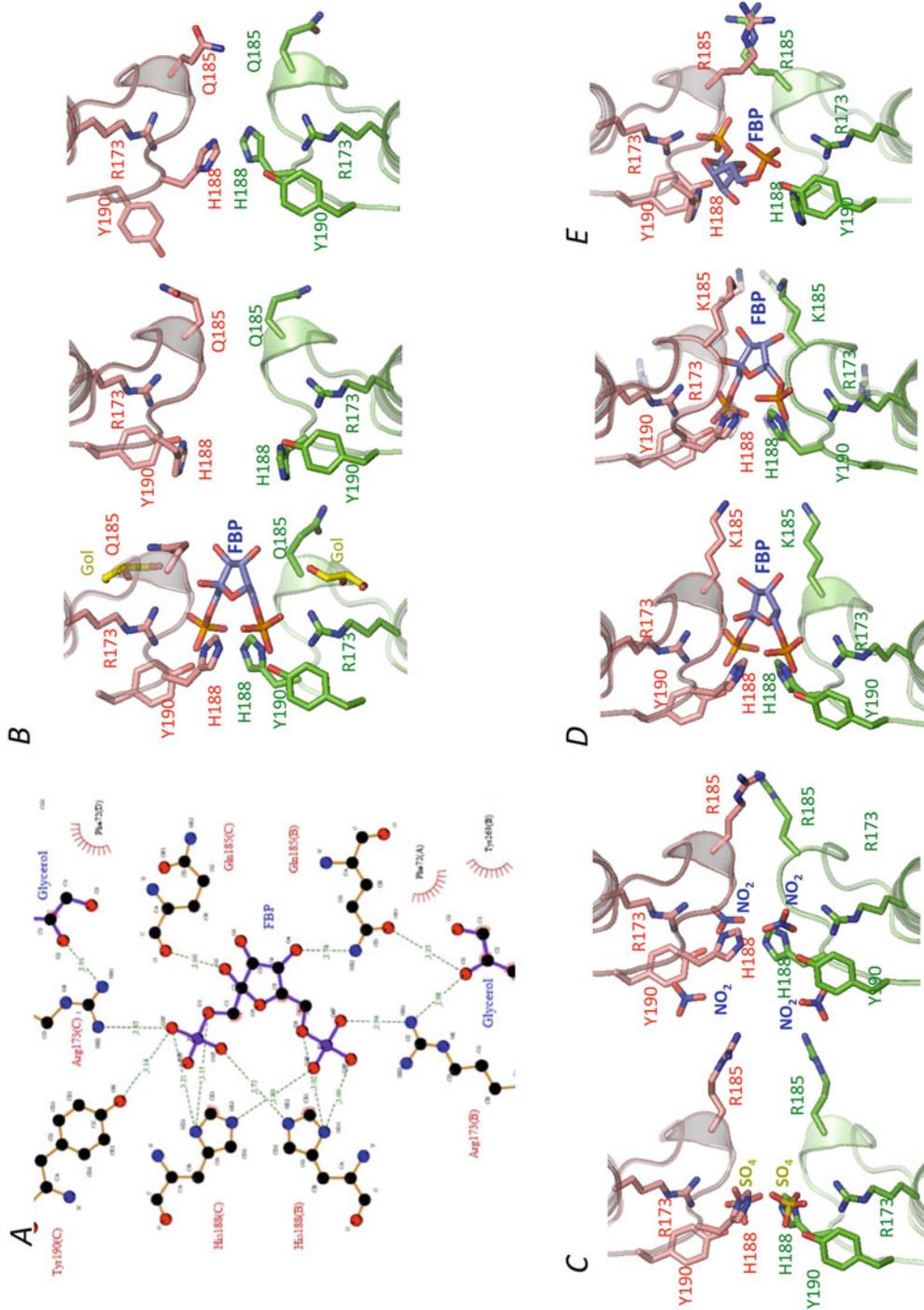
Fig. 10 Structural changes in FBP binding sites. (**a**) LIGPLOT diagram of the hydrogen bond interactions of the R state TcLDH with FBP. The A and B subunits indicated in parentheses correspond to the *red* and *green* subunits in Fig. 3a. Hydrogen bonds are indicated by *broken lines*. Numbers within the *lines* indicate the distances (Å) between the atoms corresponding to the hydrogen bonds. (**b**) the FBP binding sites of the R (*left*) and T (*center*) statea of TcLDH and R state TtLDH (*right*). The glycerol molecule in the R state TcLDH from the solvent is indicated as Gol.

(**c**) the FBP binding sites of the R (*left*) and T (*right*) state structures of LcLDH. Sulfate and nitrate from the solvent in the R and T state LcLDH structures are denoted as SO_4 and NO_3, respectively. (**d**) the FBP binding sites of the R (*left*) and T (*right*) state structures of BlLDH. (**e**) the FBP binding site of TmLDH. The *red* and *green* state subunits form contact with each other at the P-axis subunit interfaces. The two state structures are superimposed by means of least squares deviation for a P-axis relared dimer

SO$_4$ and NO$_2$ ions from the crystallization solvents, respectively (Fig. 10c). These ions apparently mimic the phosphate groups of FBP. Unlike in the case of TcLDH, the two state LcLDH structures consistently orientate the His188 imidazole to the FBP binding site, therefore His188 unlikely controls the FBP affinity, though essential for the FBP-binding. It is nevertheless notable that the area of Arg185, which locates at the MR1 of LcLDH, greatly narrows the space of the binding site in the T state structure. Hence, LcLDH likely switches the FBP affinity between the two states mostly through the MR1 motion. Furthermore, the T state BlLDH does not only orientate the His188 to the same direction as the R state enzyme, but actually binds the FBP molecule in essentially the same manner as the R state enzymes (Fig. 10d). Like Gln185 and Lys185 of TcLDH and LcLDH, Lys185 of BlLDH also narrows the space of the binding site in the T state structure. Since the MR1 motion of BlLDH is smaller than those of TcLDH and LcLDH, nevertheless, the T state BlLDH cannot completely block the FBP molecule, and therefore allows it to be accommodated in the cramped FBP binding site. In the case of BlLDH, therefore, FBP alone cannot induce the full enzyme activation, which requires the additional activation effects by substrate pyruvate. This is consistent in the catalytic properties of BlLDH, which shows a significant positive cooperativity in the pyruvate binding even in the presence of saturation level of FBP.

The TmLDH structure also contains the bound FBP molecule in the FBP-binding site, but in a quite unusual manner (Fig. 10d). In this structure, TmLDH orientates His188 outside of the binding site, and allows Arg185 to greatly narrow the space of the binding site, like the T state TcLDH, and forms much less interactions with the bound FBP molecule. as compered with the multiple interactions of the R-state TcLDH (Fig. 10a). Nevertheless, this structure suggests that TmLDH controls the affinity to FBP by a similar manner to TcLDH.

7 Structural Change in NAD-Binding Site

The NADH-binding was monitored by the change in the polarization of NADH fluorescence of NADH for TcLDH in the absence and presence of FBP (Koide et al. 1992). TcLDH exhibits no significant cooperativity on NADH binding in the absence or presence of FBP, and gives dissociation constants (K_Ds) of 0.9–1.2 μM for NADH independently of FBP. LcLDH also exhibits no significant cooperativity in the NADH-binding, which was monitored by the change in the intensity of NADH fluorescence for the T state enzyme, and gives K_D of 22 μM for NADH (Arai et al. 2010). Hence, the coenzyme binding is not directly involved in the allosteric transition in these enzymes. However, it is possible that the NADH-binding indirectly participate in the allosteric machinery, because NADH binding is virtually essential for the following pyruvate binding in the compulsory ordered bi-bi mechanism of LDH. In the case of TcLDH, the conformation of the bound NADH molecule was analyzed by means of transfer NOE with NMR spectroscopy (Koide et al. 1989, 1992; Machida et al. 1985a). The analysis indicated that the bound NADH molecule causes a change in the conformation between the T and R state structures of the enzyme, where the nicotinamide-ribose moiety of NADH takes on *syn* and *anti* conformations, respectively.

The TcLDH structures show no drastic change in the structure of NAD binding site between the two state structures (Fig. 11a). Nevertheless, the NAD-binding site is much larger than the FBP or pyruvate binding site, and therefore many small structural changes occur in the binding site between the two states: e.g. the αB and αC helices move significantly between the two states (Fig. 5). In the case of BlLDH, the bound NADH molecule consistently forms the *anti*-conformation of the nicotinamide-ribose moiety in the two state structures, although it slightly changes the orientation of the moiety between the two states (Fig. 11b). It is therefore unlikely
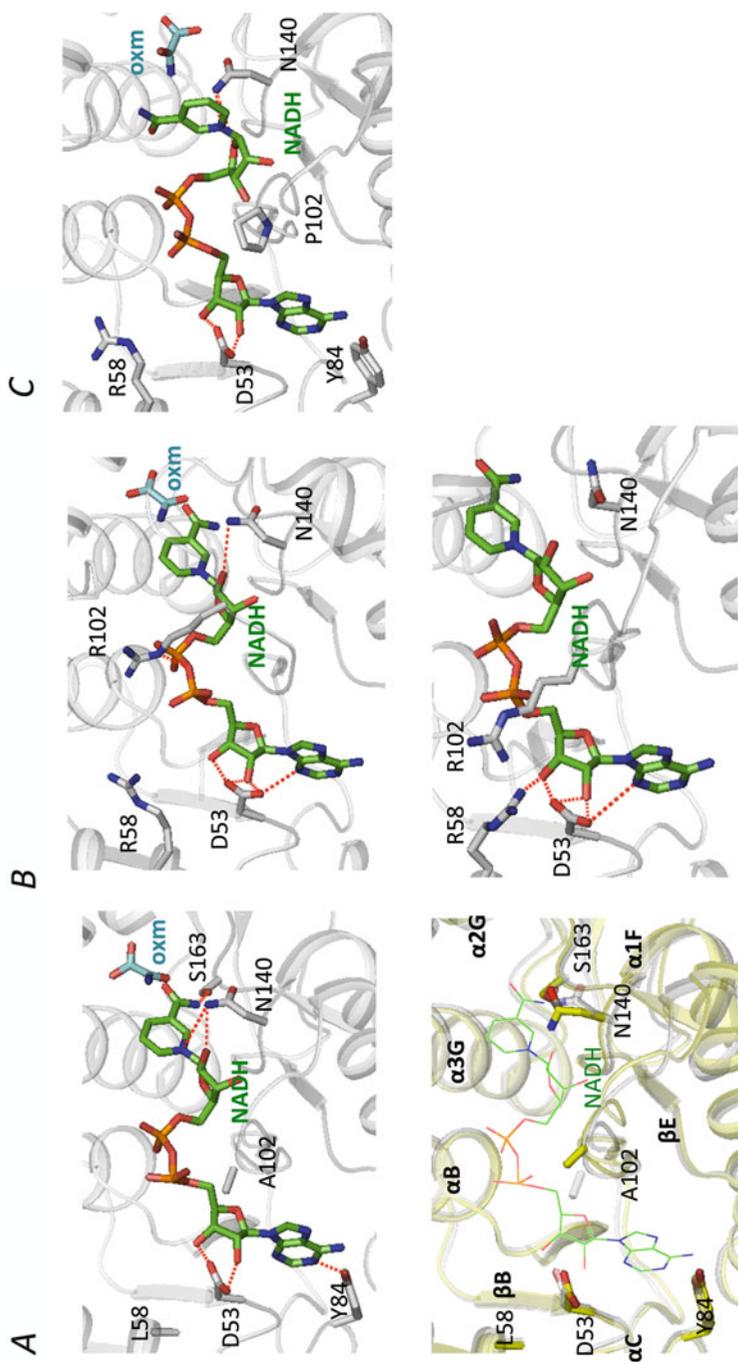
Fig. 11 Comparison of the conformations of NADH molecules bound in TcLDH, BlLDH, and TmLDH. (**a**) the structures of the NAD binding site of the **R** state (*upper*) and T state (*lower*) of TcLDH. In the *lower panel*, the **R** (*grey*) and T (*yellow*) state structures of TcLDH are superimposed. (**b**) the structures of the NAD binding site of the **R** state (*upper*) and T state (*lower*) of BlLDH. (**c**) the structures of the NAD binding site of TmLDH. *Red broken lines* indicate hydrogen bonds between NADH and the enzymes

that the regulation mediated by NADH conformation is ubiquitous in the allosteric LDHs. It is nevertheless notable that TmLDH, which is considerable to form the T state structure, contains the *syn*-conformation of the NADH nicotinamide-ribose moiety in the binding site (Fig. 11c), being consistent with the NMR study for the T state TcLDH. In the TmLDH structure, the nicotinamide- ribose of NADH forms hydrogen bonds with Asn140 in a different manner to that of the R state TcLDH or BlLDH, and shows much higher B-factor than the adenine-ribose moiety. The active site structure of TmLDH resembles that of the T state TcLDH or LcLDH much more than that of the T state BlLDH, which completely blocks the substrate binding (Fig. 9). The *syn*-conformation of NADH is unsuitable for the formation of interactions with pyruvate, and therefore likely compensates for the poor guarded active site of the T state TcLDH and TmLDH.

8   Intra-Subunit and Inter-Subunit Linkages of Allosteric Motion

The MR2 locates at the flexible surface area, where the primary and secondary structures are poorly conserved in known LDHs. In this area, TcLDH and LcLDH commonly have a α helix (the αT helix), but BlLDH does not. In the three enzymes, the MR1 and MR2 consistently form wide van der Waals contacts between them, and therefore move together in the motion of allosteric transition (Figs. 6 and 12). TcLDH forms four
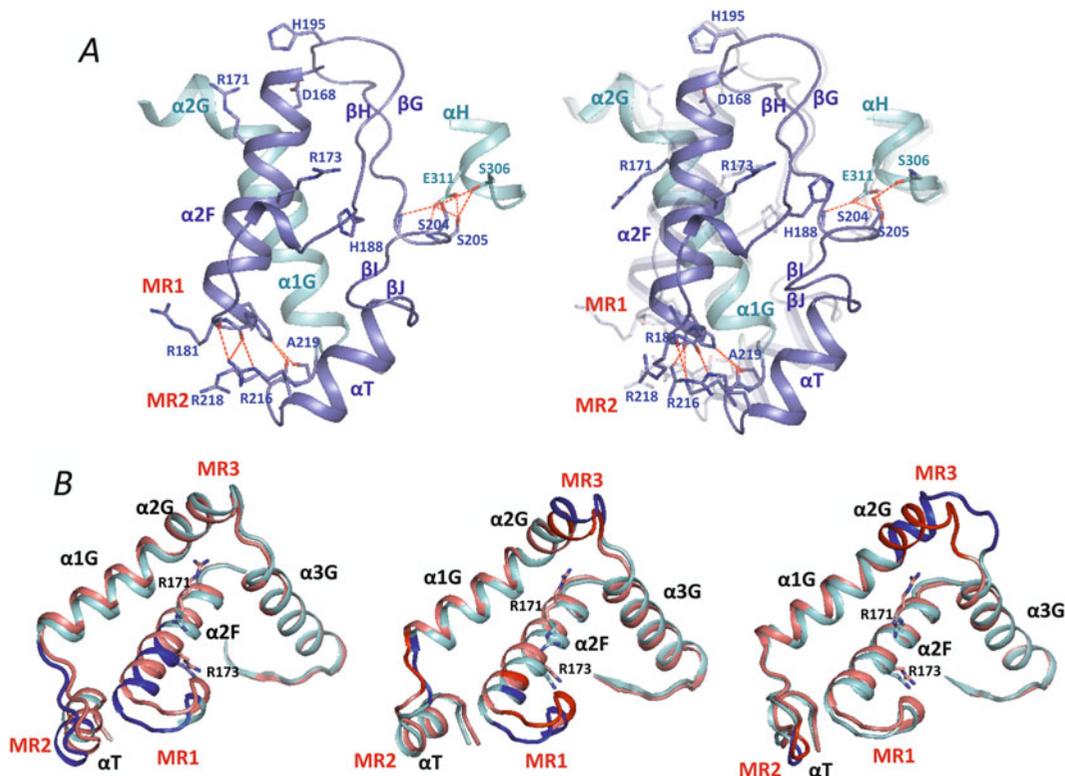
Fig. 12 Intra-subunit linkage of MR1, MR2 and MR3. (**a**) movements of MR1, MR2 and related regions in TcLDH. The R state (*left*) and T state (*right*) structure of TcLDH. In the *right panel*, the semitransparent R state structure is superimposed to the T state structure by means of least-squares deviation for Cα atoms. The *red broken lines* indicate hydrogen bonds between MR1 and MR2, and the MR1-MR2 area and the αH helix. (**b**) the intra-subunit linkages of MR1, MR2 and MR3 for TcLDH (*left*), LcLDH (*center*) and BlLDH (*right*). The R (*salmon red*) and T (*cyan*) states are superimposed by means of least-squares deviation for Cα atoms

constitutive hydrogen bonds between the two regions, and thereby constitutes a rigid mobile body of MR1 and MR2 for the allosteric motion (Fig. 12a), whereas the other two enzymes have more flexible MR1-MR2 contacts without inter-region hydrogen bond. The MR2 is highly exposed to the solvent in the opposite side to the MR1 contact, and flexibly movable with the MR1 motion. On the other hand, the MR1 also forms main contacts with the P-axis related sub-unit beside for the MR2 contact, and therefore move together with the neighboring subunit, cou-pling the tertiary structural change with the qua-ternary structural change. In the other side, the MR1 faces the cavity of LDH, which is formed in the center of tetramer, and constitutes the FBP binding site in the inner surface of the enzyme. The MR1 rotates around the N-terminal region of the α2F helix, and therefore its motion directly links to the rotation of α2F helix in the allosteric transition (Fig. 12). The motion of MR2 links to small motions of the βI-βJ sheet and the αH helix, the latter of which is highly exposed to the solvent.

In the three enzymes, the MR1-MR2 motion is also linked to the motion of the α1/2G helix (Fig. 12b). The α2F and α1/2G helices form many van der Waals contacts around Arg171 of the α2F helix and the kink of the α1/2G helix without inter-helix hydrogen bond. The α1/2G helix is highly exposed to the solvent opposite the α2F helix contact, and therefore flexibly moves depending on the α2F helix motion. On the other hand, the motion of MR3 is directly linked to the motion of the flanking α2G helix, and therefore LcLDH and BlLDH, particularly the latter enzyme greatly moves the α2G part of α1/2G helix, whereas TcLDH moves the α2G helix only slightly. Hence, the flexible α1/2G helix adsorbs both the MR1-MR2 and MR3 motion with the kink of the helix, and cancels the intra-subunit linkage between them in the three enzymes.

TcLDH and BlLDH apparently allow the motions of MR1-MR2 and MR3 to be coupled with each other only through the quaternary structural changes of the enzymes. Particularly TcLDH undergoes a uniquely simple structural change in the allosteric transition, using the MR1-MR2 as the core for the allosteric motion. TcLDH only couples the MR1-MR2 motion with the quaternary structural change at the P-axis subunit interface (Fig. 7a), and cancels the cou-pling at the Q-axis subunit interface using flexi-ble motions of the αB and αC helices (Fig. 8d). On the other hand, LcLDH apparently forms an additional pathway for the linkage motion between MR1-MR2 and MR3, through the motion of the α3G helix. In LcLDH, the MR1 motion does not only link to the quaternary struc-tural change, but also links to the specific motion of the α3G helix of the neighboring subunit (Figs. 7b), and the α3G motion along the helical axis, which is supported by flexible motion of the αB(Q) helix, links to the motion of the α2G-α3G loop area (MR3) (Figs. 8e). In the P-axis related dimer of LcLDH, hence, two sets of the linkage forms circular and symmetrical frame of the linkage around the P-axis, comprising two sub-strate binding sites at the two apices and one FBP binding site at the center of the frame (Fig. 13). This linkage consists of the tight junctions of inter-subunit contacts (pale pink areas in Fig. 13) between the α2F-βG and α3G-βK (P) loops at the P-axis subunit interface (Fig. 7b), and between the α2G-α3G loop and αC(P) helix at the Q-axis interface (Fig. 8b), and its flexibility for the allosteric motion is endorsed by the flexible contacts (pale yellow areas) between the α2F and αC(Q) helices (Figs. 9c and 14b). The FBP molecule only fits and fixes the R state circular frame at the center. The allosteric motion of LcLDH by the large mobile frame is in contrast to the motion of TcLDH by the mobile structural core.

9 Enzyme Modifications Altering the Allosteric Properties

Whereas the vertebrate LDHs apparently fix the R state (open) conformations with the N-terminal R-arm sequences, bacterial non-allosteric LDHs lack the R arm sequence as in the case of the allosteric enzymes (Figs. 1 and 3). Nevertheless, *Lactobacillus pentosus* LDH (LpLDH)
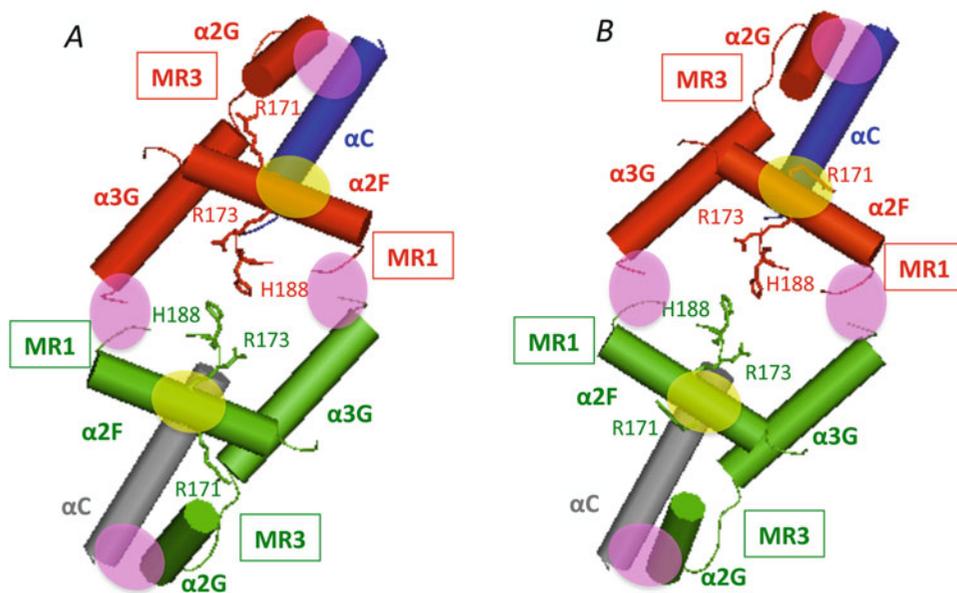
Fig. 13 Intersubunit linkages for the allosteric motion in the P-axis related dimer of LcLDH. The secondary structural elements of four different subunits (*red*, *green*, *blue*, and *grey*) are viewed from P-axis for the R (**a**) and T (**b**) state structures of LcLDH. *Circles in pale pink* indicate tight intersubunit contact areas in the linkage that contain constitutive inter-subunit salt bridges and hydrogen bonds (Figs. 7b and 8b). *Circles in pale yellow* indicate flexible intersubunit contact areas that contain no intersubunit salt bridge or hydrogen bond (Figs. 9c and 14b)

constitutively exhibits high catalytic activity independently of FBP in spite of its particularly high amino acid sequence identity (67 %) with LcLDH (Taguchi and Ohta 1992), implying that the exchange between allostery and non-allostery can readily occur in LDHs through only small structural changes, such as some amino acid replacements. LpLDH possesses a unique inter-subunit salt bridge network in the α2F-αC(Q) contact area, where basic residues, Arg171 and Lys178, in the α2F helix, and Lys235, in the α2G helix, form multiple inter-subunit salt bridges with acidic residues, Asp67 and Glu68, in the αC helix of the counterpart subunit (Fig. 14a) (Uchikoba et al. 2002). This salt bridge network appears to tightly lock the α2F and αC(Q) helices, together with the orientation of Arg171. In contrast, LcLDH does not possess these amino acid residues beside for Arg171, and therefore appears to flexibly change the structure of the corresponding inter-subunit region (Fig. 14b). The mutant LcLDH (Q5-LcLDH), which mimics the salt-bridge network of LpLDH by replacements of

five amino acids, S67E/N68D/E178K/A235K/D234N is designed to evaluate the exact role of the salt-bridge network (Arai et al. 2011). LpLDH-like salt-bridge networks are found also in many non-allosteric LDHs, or L-malate dehydrogenases (MDHs) that have the LDH-like tertiary structure and the Q-axis related dimeric structures. *Thermus thermophilus* L-malate dehydrogenase (TtMDH) has virtually the same network to that of LpLDH (Fig. 14c). Although TtMDH has Arg235 instead of Lys235, Arg235 may rather strengthen the network by its guanidinium side chain than Lys235. In contrast, the allosteric LDHs including TcLDH and BlLDH consistently lack the salt-bridge network (Fig. 14d and e). In TcLDH, the mutant enzymes mimicking LpLDH (Q4K, L67E/H68D/E178K/A235K) and TtLDH (Q4R, L67E/H68D/E178K/A235R) are designed (Ikehara et al. 2014).

The Q5 mutant LcLDH exhibits high catalytic activity with small substrate K_m independently of FBP, although it still shows the significant FBP-dependence at pH 7.0 (Fig. 15a) (Arai et al. 2011). On the other hand, the Q4R
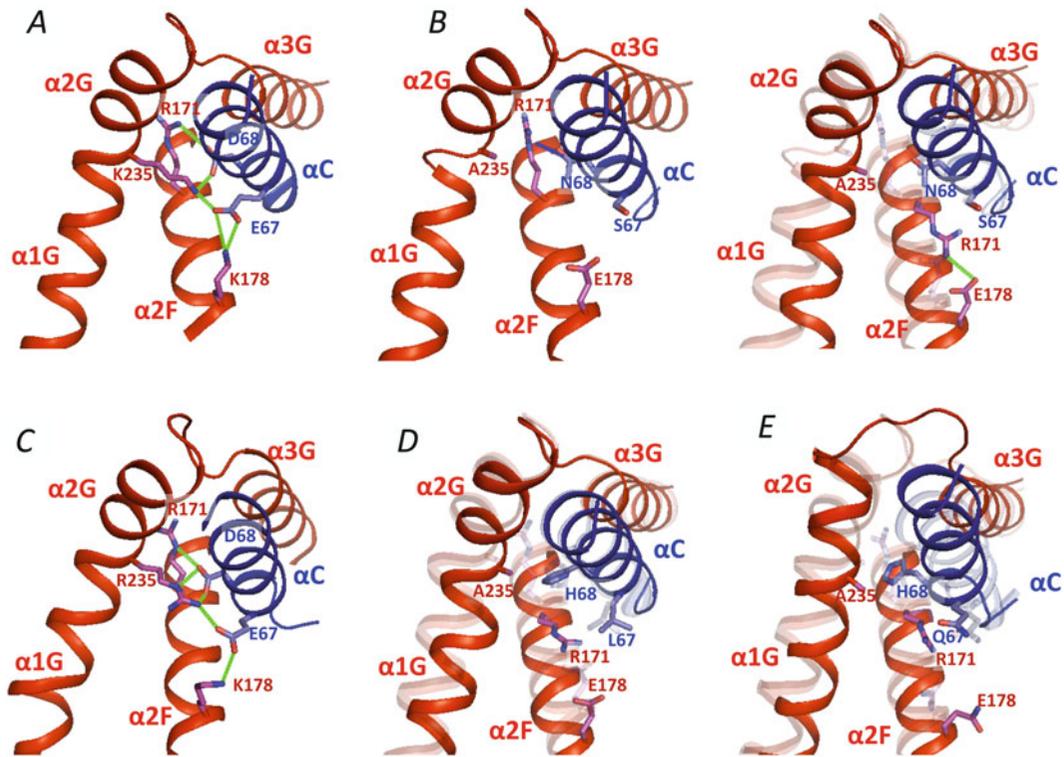
Fig. 14 The Q-axis inter-subunit regions around α2F and αC(Q) helices. (**a**) LpLDH. (**b**) the R state structure (*left*) and the T state structure are superimposed by means of least squares deviation with the R state structure (semi-transparent) (*right*) of LcLDH. (**c**) *Thermus thermophilus* L-malate dehydrogenase (PDB number 1IZ9). (**d** and **e**), superimpositionings of T state and R state (*semitransparent*) structures of TcLDH (**a**) and BlLDH (**a**). The structures are represented by *ribbon diagrams*, where only the amino acids at positions 67, 68, 171, 178 and 235 are indicated by a *stick model*. The two Q-axis-related subunits are colored *red* and *blue*. Salt bridges and hydrogen bonds among these amino acids are indicated by *green* and *blue broken lines*, respectively

TcLDH shows a great FBP-dependence as in the case of the wild-type TcLDH, although it exhibits markedly improved substrate K_m and V_{max} in the absence of FBP (Fig. 15b) (Ikehara et al. 2014). The Q4K mutant TcLDH shows essentially the same catalytic profile as the Q4R enzyme. The different effects of virtually the same mutation are explained by the followings. First of all, the change in Gibbs free energy (ΔG) between the two states, which is calculated by kinetic analysis (Koide et al. 1992) or binding analysis for oxamate (Arai et al. 2010; Iwata et al. 1994), is estimated to be 7.3 kcal/mol for TcLDH (Koide et al. 1992), 8.0 kcal/mol for BlLDH (Iwata et al. 1994) and 3.4 kcal/mol for LcLDH (at pH 5.5) (Arai et al. 2010). The Q4 mutations may not overcome the greater ΔG of

TcLDH. In second, LcLDH exhibits larger structural changes in the corresponding Q-axis interface in the allosteric transition, as compared with TcLDH (Figs. 8, 13 and 14). The Q-axis network forms salt-bridges between the α2F and αC (Q) helices, and between the α2G and αC (Q) helices. Therefore, the network is effective for LcLDH, which greatly moves both the α2F and αC(Q) helices together with α2G-α3G area, more than TcLDH, which moves only the α2F helix. LcLDH possesses the large mobile frame for the allosteric motion (Fig. 13), and the Q-axis salt-bridge network may effectively fix the R state frame, through fixing the flexible contacts in the frame. In addition, LcLDH is closely relative to LpLDH, sharing 67 % of identical amino acid residue. LpLDH possesses virtually the
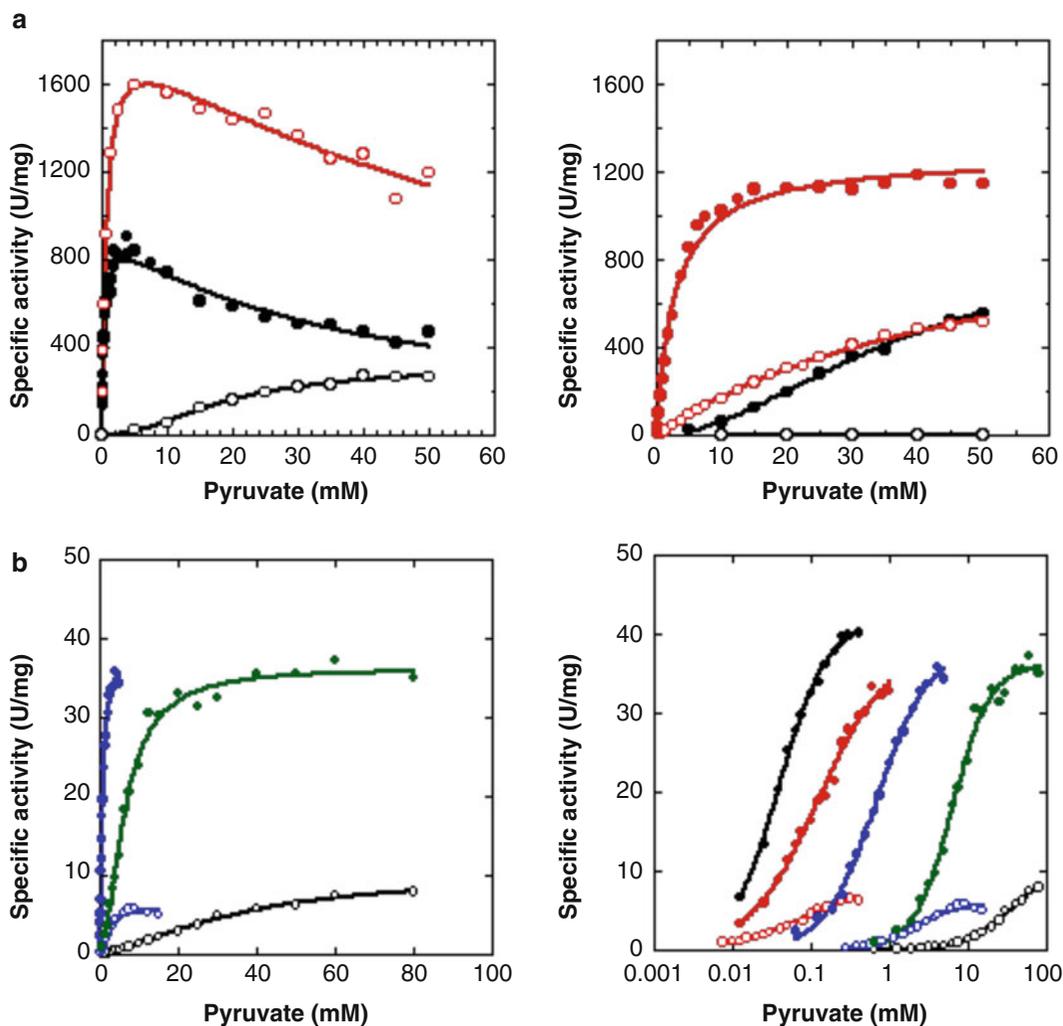
Fig. 15 Catalytic profiles of mutant LcLDH and TcLDH at 30 °C. (**a**) pyruvate saturation profiles for wild-type (*black*) and Q5 mutant (*red*) LcLDHs at pH 5.0 (*left*) and pH 7.0 (*right*) without FBP (*open symbols*) and with 5 mM FBP (*closed symbols*). (**b**) pyruvate saturation curves for the wild-type TcLDH (*black*) with (*closed symbols*) and without (*open symbols*) 0.1 mM FBP, and for the Q4 (*closed green symbols*), R173Q (*open blue symbols*), Q4/R173Q (*closed blue symbols*), R173Q/R216L (*open red symbols*) and Q4/R173Q/R216L (*closed red symbols*) mutant TcLDHs without FBP. The concentrations of pyruvate are plotted on linear (*left*) and logarithmic (*right*) scales

same frame as LcLDH, forming inter-subunit contact area with conserved amino acids. The frame of LpLDH only differs from that of LcLDH in the α2F-αC(Q) contacts, in which LpLDH forms the tight contact with the salt-bridge network instead of the flexible contact of LcLDH. In contrast, TcLDH only moves the α2F helix in the area of the salt-bridge network (Figs. 9 and 14) using the mobile core of

MR1-MR2 (Fig. 12), which is distant from the corresponding area. Since of the smaller structural change, the Q4 mutation likely less selectively stabilizes both the R and T state structures in TcLDH.

On the other hand, TcLDH is uniquely activated through chemical modifications for its Arg residues with 2,3-butanedione under protection of the catalytic site with NADH and oxamate

(Taguchi et al. 1984). Such a phenomenon has not been reported in other LDHs, beside for *Thermus aquaticus* LDH (Machida et al. 1985b), which is closely related to TcLDH. Transfer NOE with NMR spectroscopy strongly suggests that the 2,3-butanedione-modified TcLDH forms the fully active structure (Koide et al. 1989, 1992). The FBP-independent activity of TcLDH is also markedly enhanced by the replacement of Arg173 (R173Q) and Arg216 (R216L), which are predominantly modified with 2,3-butanedione (Koide et al. 1992; Matsuzawa et al. 1988). These amino acid replacements exhibit additive activation effects on the Q4 mutant enzymes (Fig. 15b) (Ikehara et al. 2014), although these replacements do not completely activate TcLDH, unlike FBP or 2,3-butandione (Koide et al. 1989 1992; Machida et al. 1985a).

10 Mediation of Allosteric Equilibrium by Repulsive Static Interactions

TcLDH, LcLDH and BlLDH consistently undergo their allosteric transitions according to the MWC theory, greatly biasing the allosteric (pre-existing) equilibrium to the T state structure. Although the T state structure should be therefore stable more than the R state structure, the two state structures of these enzymes do not appear to clearly show the apparent factor to stabilize the T state structure. The T state conformations are apparently not more compact than the R state structures in the three enzymes. In addition, the three enzymes only form comparative numbers of inter-subunit hydrogen bond or salt-bridge in the two state structures (Table 1). It is interesting that non-allosteric LpLDH forms markedly increased numbers of inter-subunit hydrogen bonds and salt-bridges, implying that allosteric LDHs limit the inter-subunit hydrogen bonds for their flexible motions of quaternary structure even in the heat stable enzymes such as TcLDH. It is therefore rather possible that T state structure is only relatively stabilized by the factor that destabilizes the R state structure in the

allosteric equilibrium in the cases of these three enzymes. From this point of view, it is interesting that the modifications for basic amino acid residues markedly increase the FBP-independent activity of TcLDH. These phenomenons strongly suggest that the positively charged groups within the protein molecule destabilize the R state structure, and consequently increase the relative stability of the T state structure in TcLDH. Unsurprisingly, the factor considerable should be the repulsive Coulomb's force within the protein molecule.

In the surface of TcLDH structure, patches of positive surface potential exist around the R-symmetry axis, and are focused on and disperse from the axis in the R and T states, respectively (Fig. 16a and b). The main positive patches near the R axis are generated by Met22 (the N-terminal amino group), Arg47, His74 and Arg92, which form a cluster of positive charges around the core β-sheet of the Rossmann-fold NAD-binding domain, together with Lys23 and Arg82 (named the N-domain cluster). These patches faced in the R-axis approach closely to each other in the R state structure, and separate from each other in the T state structure. Therefore, these clusters must predominantly destabilize the R structure (closed conformation) though strong static repulsions between the R-axis (also the Q-axis) related subunits. It is further interesting that the other main patches, which locate vicinity of the N-domain cluster, are generated mostly from Arg181 and Arg218, which locate at MR1 and MR2, respectively. These residues form a cluster of positive charges together with Arg173, His179 and Arg216 in the MR1-MR2 area (named the MR1-MR2 cluster) in the outer and inner surface of TcLDH. In the R state structure, the MR1-MR2 clusters do not only approach to each other, but also approach closely to the N-domain clusters of the Q-axis (also the R-axis) related subunits. While the N-domain cluster only moves together with the quaternary structural change, the MR1-MR2 clusters move in association to both the tertiary and quaternary structural changes, and therefore more greatly change the distances to the counterpart clusters between the two state structures (Fig. 16c). In the
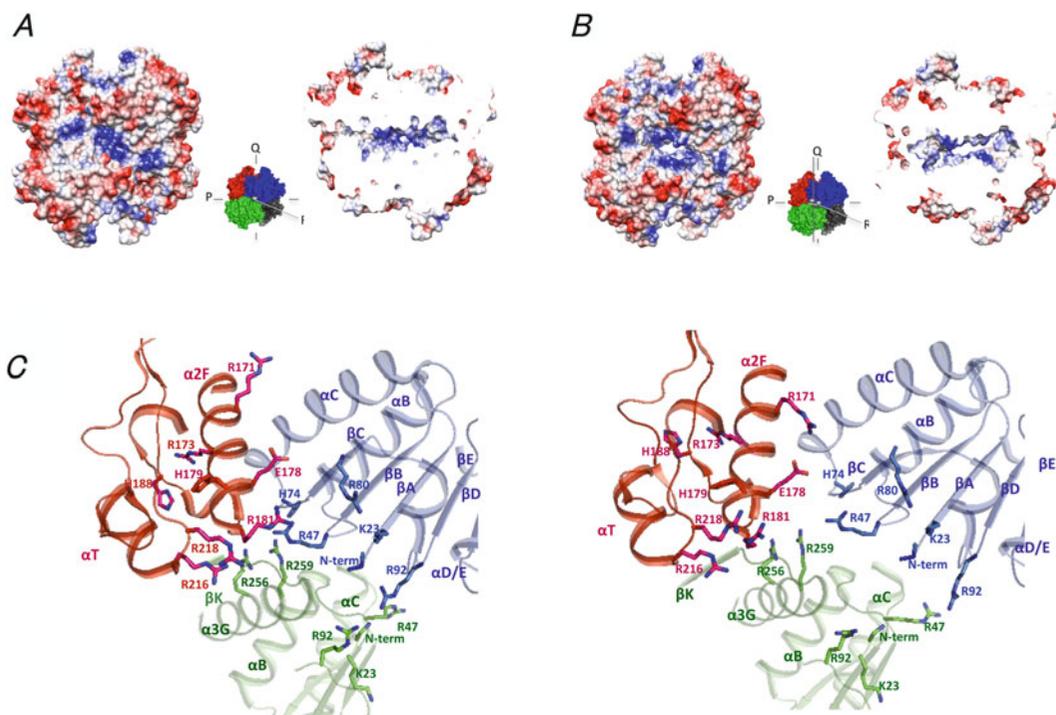
Fig. 16 Clusters of positive charges and allosteric motion of TcLDH. (**a**) and (**b**), surface electrostatic potential models of the R (**a**) and T (**b**) state structures of TcLDH in outer surface (*left*) and inner surface (*right*) of the tetramer. The models show electronegative (*red*), electropositive (*blue*), and electroneutral (*white*) amino acid side chains. (**c**) structures of the cluster areas of positively charged residues in the R (*left*) and T (*right*) states of TcLDH. The subunits are colored according to panel (**a**)

inner surface of the enzyme, Arg256 and Arg259 form additional clusters, which also approach to and separate from each other and the N-domain cluster in the R and T state structures, respectively. This cluster moves only through the quaternary structural change, as in the case of the N-domain cluster. These clusters consistently change the distance to counterpart clusters only between the Q and R-axis related subunits, but do not markedly change between the P-axis related dimer.

The observation in the surface electric potential clearly explains the apparent enzyme activation by R173Q and R216L replacements, which reduce the positive charges of the MR1-MR2 cluster, and eliminate the static repulsions between the positive charged groups. This also explains the apparently higher FBP-independent activity (smaller substrate K_m) of TtLDH, which has Tyr179 instead of His179. The H179Y

replacement does not only reduce the positive charges, but also relaxes the MR1-MR2 linkage, in which His179 form pivotal inter-region hydrogen bonds.

Naturally, the change of allostery markedly affects the stability of protein (Fig. 17). The wild-type TcLDH markedly reduce the apparent substrate $S_{0.5}$ values as increase of temperature in the absence of FBP (Fig.17a), while it shows no marked temperature dependency in the substrate K_m value, suggesting that the allosteric equilibrium markedly depends on temperature (Ikehara et al. 2014). On the other hand, the Q4R mutant enzyme shows less temperature dependency in the $S_{0.5}$ than the wild-type enzyme. It is also true with TtLDH (Colletier et al. 2012). This suggests that the Q4R mutation change the relative thermal stabilities of the two state structures. In LcLDH, the Q5 mutation does not increase, or even reduces the thermal stability of the enzyme
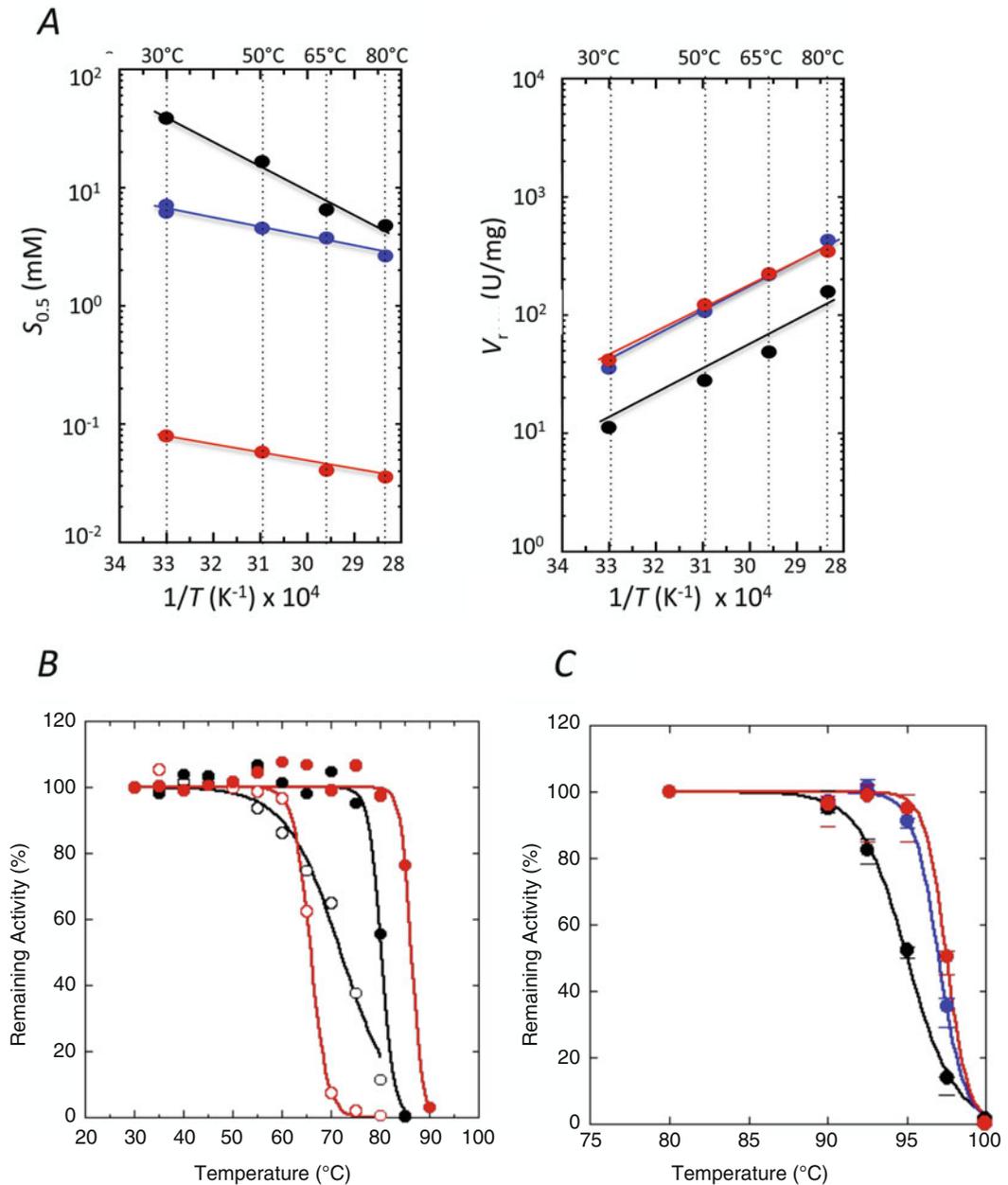
Fig. 17 Thermal profiles of the wild-type, and Q4 and Q4/P2 mutant TcLDHs. (**a**) Arrhenius plots of the $S_{0.5}$ (*left*) and V_{max} (*right*) values for the wild-type (*black*), and Q4 (*blue*) and Q4/R173Q/R216L (*red*) mutant TcLDHs. (**b**) Thermal stability of the wild-type (*black*) and Q5-mutant LcLDH (*red*) with (*closed symbols*) and without (*open symbols*) 5 mM FBP.

(**c**) Thermal stability of the wild-type (*black*), Q4 (*blue*) and Q4/R173Q/R216L (*red*) mutant TcLDH. The enzymes were diluted to 0.3 mg/ml with 100 mM sodium acetate buffer (pH 5.0) for the LcLDHs, and with 50 mM sodium MOPS buffer (pH 7.0) for the TcLDHs, and then treated for 30 min at the indicated temperatures

(Fig.17b), although the mutation introduces inter-subunit hydrogen bonds to the enzyme, increasing the activation enthalpy ($\Delta H^{\ddagger}$) and activation entropy ($\Delta S^{\ddagger}$) of the heat-inactivation

(Arai et al. 2011). This suggests that the Q5 mutation does not only stabilize the R state structure, but also destabilizes the T state structure of LcLDH. In the presence of FBP, unsurprisingly,

the mutant LcLDH exhibits markedly increased thermal stability, being consistent in the MWC model, in which the allosteric activator specifically stabilizes the R state structure. On the other hand, TcLDH is a highly thermal stable enzyme from an extreme thermophile (Taguchi et al. 1982). The Q4R mutation significantly increases the thermal stability of TcLDH (Fig.17c) (Ikehara et al. 2014). This also indicates the Q4R mutation likely stabilizes both the T and R state structure unselectively in TcLDH, unlike the Q5 mutation in LcLDH. It is further notable that R173Q/R216 mutation also increases the thermal stability of the enzyme. Nevertheless, this may not be very surprising, since the static repulsion within protein molecule is generally a negative factor for protein stability. This mutation likely eliminates such a negative factor even from the T state structure of TcLDH, although the static repulsion predominantly destabilizes the R state structure. It may be more noteworthy that even thermophilic LDHs such as TcLDH employ such a negative factor as static repulsion for the allosteric regulation.

Although LcLDH or BlLDH does not show such a simple surface profile as TcLDH, it is possible that many allosteric LDHs including the two enzymes also use electrostatic repulsions to balance their allosteric equilibrium. The bacterial LDHs can readily change the surface electrostatic potential by only a few amino acid replacements, and therefore readily change and diversify their allosteric properties during evolution. *Geobacillus stearothermophilus* LDH (GsLDH), which is a well studies bacterial allosteric LDH (Clarke et al. 1989), likely uses the electrostatic repulsion for the allosteric regulation in somewhat different manner from TcLDH. GsLDH cannot retain the stable tetrameric structure unless FBP is present (Clarke et al. 1985, 1987). In the absence of FBP, the GsLDH tetramer is readily dissociated into the stable dimer (Q-axis related dimer) at physiological concentration. The dimeric GsLDH has a T state-like tertiary structure, in which Arg171 is orientated outside the active site, and exhibits low catalytic activity with large substrate K_m values (Cameron et al. 1994). Therefore, GsLDH cannot induce

the allosteric transition by pyruvate alone, and consistently shows hyperbolic shapes of the pyruvate saturation curve. In GsLDH, the R173Q mutation induces the desensitization to FBP without apparent enzyme activation, and stabilizes the tetrameric structure in the absence of FBP (Clarke et al. 1987). In the R173Q mutant GsLDH, FBP rather destabilizes the tetrameric structure, suggesting that the tetrameric structure of GsLDH is destabilized mostly by static repulsion at the p-axis subunit interface. GsLDH does not possess the stable tetrameric T state structure that can bear static repulsions within protein, instead possesses the stable T state dimeric structure for the allosteric regulation.

11 Conclusion

TcLDH, LcLDH and BlLDH consistently undergo the allosteric transition according to the Monod-Wyman-Changeux (pre-existing equilibrium) model, and take on open and closed conformations of homotetramers for the T and R states, coupling the quaternary structural changes, the tertiary structural changes, and the structural changes in the binding sites for substrate and FBP. Nevertheless, the three enzymes exhibit markedly different structural changes from one another in the tertiary structures and the ligand binding sites, indicating a high variety in the allosteric machineries of bacterial LDHs. LcLDH and BlLDH greatly change the structures of both the P and Q-axis subunit interfaces in their allosteric motions. Particularly, LcLDH apparently uses large mobile frames, which form inter-subunit linkages for the allosteric motion. In contrast, TcLDH exhibits the simplest allosteric motion in the three enzymes, involving a simple mobile structural core in its allosteric motion, which couples the tertiary and quaternary structural changes at the P-axis subunit interface. The uniquely simple allosteric motion and surface electric potential of TcLDH indicate that the enzyme mediates the allosteric equilibrium mostly through electrostatic repulsion within the protein molecule.

The repulsive electrostatic interactions are likely involved in the regulations of many bacterial allosteric LDHs.

References

Abad-Zapetero C, Griffith JP, Sussman JL, Rossmann MG (1987) Refined crystal structure of dogfish M4 apo-lactate dehydrogenase. J Mol Biol 198:445–467

Arai K, Ishimitsu T, Fushinobu S, Uchikoba H, Matsuzawa H, Taguchi H (2010) Active and inactive state structures of unliganded *Lactobacillus casei* allosteric L-lactate dehydrogenase. Proteins 78:681–694

Arai K, Ichikawa J, Nonaka S, Miyanaga A, Uchikoba H, Fushinobu S, Taguchi H (2011) A molecular design that stabilizes active state in bacterial allosteric L-lactate dehydrogenases. J Biochem 150:579–591

Auerbach G, Ostendorp R, Prade L, Kornörfer I, Dams T, Huber R, Jaenicke R (1998) Lactate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*: the structure at 2.1 Å resolution reveal strategies for intrinsic protein stabilization. Structure 6:769–781

Bur D, Clarke T, Friesen JD, Gold M, Hart KW, Holbrook JJ, Jones JB, Luyten MA, Wilks HM (1989) On the effect on specificity of Thr246-Gly mutation in L-lactate dehydrogenase of *Bacillus sterothermophilus*. Biochem Biophys Res Commun 161:59–63

Cameron AD, Roper DI, Moreton KM, Mirhead H, Holbrook JJ, Wigley DB (1994) Allosteric activation in *Bacillus stearothermophilus* lactate dehydrogenase investigated by an X-ray crystallographic analysis of a mutant designed to prevent tetramerization of the enzyme. J Mol Biol 238:615–625

Clarke AR, Atkinson T, Campbell JW, Holbrook JJ (1985) The assembly mechanism of the lactate dehydrogenase tetramer from *Bacillus stearothermophilus*; the equilibrium relationships between quaternary structure and the binding of fructose 1,6-biphosphate,

NADH and oxamate. Biochim Biophys Acta 829:387–396

Clarke AR, Wigley DB, Chia WN, Barstow DA, Atkinson T, Holbrook JJ (1986) Site-directed mutagenesis reveals role of mobile arginine residue in lactate dehydrogenase catalysis. Nature 324:699–702

Clarke AR, Wigley DB, Barstow DA, Chia WN, Atkinson T, Holbrook JJ (1987) A single amino acid substitution deregulates a bacterial lactate dehydrogenase and stabilizes its tetrameric structure. Biochim Biophys Acta 913:72–80

Clarke AR, Wilks HM, Barstow DA, Atkinson T, Chia WN, Holbrook JJ (1988) An investigation of the contribution made by the carboxylate group of an active site histidine-aspartate couple to binding and catalysis in lactate dehydrogenase. Biochemistry 27:1617–1622

Clarke AR, Atkinson T, Holbrook JJ (1989) From analysis to synthesis: new ligand binding sites on the lactate dehydrogenase framework. Trends Biochem Sci 14:101–105, 145–148

Colletier JP, Aleksandrov A, Coquelle N, Mraihi S, Mendoza-Barberá E, Field M, Madern D (2012) Sampling the conformational energy landscape of a hyperthermophilic protein by engineering key substitutions. Mol Biol Evol 29:1683–1694

Coquelle N, Fioravanti E, Weik M, Vellieux F, Madern D (2007) Activity, stability and structural studies of lactate dehydrogenases adapted to extreme thermal environments. J Mol Biol 374:547–562

Eszes CM, Sessions RB, Clarke AR, Moreton KM, Holbrook JJ (1996) Removal of substrate inhibition in a lactate dehydrogenase from human muscle by a single residue change. FEBS Lett 399:193–197

Eventoff W, Rossmann MG, Taylor SS, Torff HJ, Meyer H, Keil W, Kiltz HH (1977) Structural adaptations of lactate dehydrogenase isozymes. Proc Natl Acad Sci U S A 74:2677–2681

Fischer E (1894) Einfluss der Configuration auf die Wirkung der Enzyme. Ber Dt Chem Ges 27:2985–2993

Fushinobu S, Kamata K, Iwata S, Sakai H, Ohta T, Matsuzawa H (1996) Allosteric activation of L-lactate dehydrogenase analyzed by hybrid enzymes with effector-sensitive and -insensitive subunits. J Biol Chem 271:25611–25616

Fushinobu S, Ohta T, Matsuzawa H (1998) Homotropic activation via the subunit interaction and allosteric symmetry revealed on analysis of hybrid enzymes of L-lactate dehydrogenase. J Biol Chem 273:2971–2976

Garvie EI (1980) Bacterial lactate dehydrogenases. Microbiol Rev 43:106–139

Grau UM, Trommer WE, Rossmann MG (1981) Structure of the active ternary complex of pig heart lactate dehydrogenase with S-lac-NAD at 2.7 A resolution. J Mol Biol 151:289–307

Hart KW, Clarke AR, Wigley DB, Waldman ADB, Chia WN, Barstow DA, Atkinson T, Jones JB, Holbrook JJ (1987) A strong carboxylate-arginine interaction is important in substrate orientation and recognition in

lactate dehydrogenase. Biochim Biophys Acta 914:294–298

Holbrook JJ, Liljas A, Steindel SJ, Rossmann MG (1975) Lactate dehydrogenase. In: Boyer PD (ed) The enzymes, vol 11, 3rd edn. Academic, New York, pp 191–292

Ikehara Y, Arai K, Furukawa N, Ohno T, Miyake T, Fushinobu S, Nakajima M, Miyanaga A, Taguchi H (2014) The core of allosteric motion in *Thermus caldophilus* L-Lactate dehydrogenase. J Biol Chem 289:31550–31564

Iwata S, Ohta T (1993) Molecular basis of allosteric activation of bacterial L-lactate dehydrogenase. J Mol Biol 230:21–27

Iwata S, Kamata K, Minowa T, Ohta T (1994) T and R states in the crystals of bacterial L-lactate dehydrogenase reveal the mechanism for allosteric control. Nat Struct Biol 1:176–185

Koide S, Yokoyama S, Matsuzawa H, Miyazawa T, Ohta T (1989) Conformation of NAD$^+$ bound to allosteric L-lactate dehydrogenase activated by chemical modification. J Biol Chem 264:8676–8679

Koide S, Yokoyama S, Matsuzawa H, Miyazawa T, Ohta T (1992) Conformational equilibrium of an enzyme catalytic site in the allosteric transition. Biochemistry 31:5362–5368

Koshland DE (1958) Application of a theory of enzyme specificity to protein synthesis. Proc Natl Acad Sci U S A 44:98–104

Koshland DE, Némethy G, Filmer D (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry 5:365–368

Machida M, Yokoyama S, Matsuzawa H, Miyazawa T, Ohta T (1985a) Allosteric effect of fructose 1,6-bisphosphate on the conformation of NAD$^+$ as bound to L-lactate dehydrogenase from *Thermus caldophilus* GK24. J Biol Chem 260:16143–16147

Machida M, Matsuzawa H, Ohta T (1985b) Fructose 1,6-bisphosphate-dependent L-lactate dehydrogenase from *Thermus aquaticuas* YT-1, an extreme thermophilile: activation by citrate and modification reagents and comparison with *Thermus caldophilus* GK24 L-lactate dehydrogenase. J Biochem 97:899–909

Matsuzawa H, Machida M, Kunai K, Ito Y, Ohta T (1988) Identification of an allosteric site residue of a fructose 1,6-bisphosphate-dependemt L-lactate dehydrogenase of *Thermus caldophilus* GK24: production of a non-allosteric form by protein engineering. FEBS Lett 233:375–378

Monod J, Wyman J, Changeux JP (1965) On the nature of allosteric transitions: a plausible model. J Mol Biol 12:88–118

Motlagh HN, Wrabl JO, Li J, Hilser VJ (2014) The ensemble nature of allostery. Nature 508:331–339

Sakowicz R, Kallwass HK, Parris W, Kay CM, Jones JB, Gold M (1993) Threonine 246 at the active site of the L-lactate dehydrogenase of *Bacillus stearothermophilus* is important for catalysis but not for substrate binding. Biochemistry 32:12730–12735

Taguchi H, Ohta T (1992) Unusual amino acid substitution in the anion-binding site of *Lactobacillus plantarum* non-allosteric L-lactate dehydrogenase. Eur J Biochem 209:993–998

Taguchi H, Yamashita M, Matsuzawa H, Ohta T (1982) Heat-stable and fructose 1,6-bisphosphate-activated L-lactate dehydrogenase from an extremely thermophilic bacterium. J Biochem 91:1345–1348

Taguchi H, Matsuzawa H, Ohta T (1984) L-Lactate dehydrogenase from *Thermus caldophilus* GK24, an extremely thermophilic bacterium. Desensitization to fructose 1,6-bisphosphate in the activated state by arginine-specific chemical modification and the N-terminal amino acid sequence. Eur J Biochem 145:283–290

Taguchi H, Machida M, Matsuzawa H, Ohta T (1985) Allosteric and kinetic properties of L-lactate dehydrogenase from *Thermus caldophilus* GK24, an extremely thermophilic bacterium. Agric Biol Chem 49:359–364

Uchikoba H, Fushinobu S, Wakagi T, Konno M, Taguchi H, Matsuzawa H (2002) Crystal structure of non-allosteric L-lactate dehydrogenase from *Lactobacillus pentosus* at 2.3 Å resolution: specific interactions as subunit interfaces. Proteins 46:206–214

Wigley DB, Gamblin SJ, Turkenburg JP, Dodson EJ, Piontek K, Muirhead H, Holbrook JJ (1992) Structure of a ternary complex of an allosteric lactate dehydrogenase from *Bacillus stearothermophilus* at 2.5 Å resolution. J Mol Biol 223:317–335

Reduction of Chemically Stable Multibonds: Nitrogenase-Like Biosynthesis of Tetrapyrroles

Gunhild Layer, Joern Krausze, and Jürgen Moser

Abstract

The sophisticated biochemistry of nitrogenase plays a fundamental role for the biosynthesis of tetrapyrrole molecules, acting as key components of photosynthesis and methanogenesis. Three nitrogenase-like metalloenzymes have been characterized to date. Synthesis of chlorophylls and bacteriochlorophylls involves the reduction of the C17-C18 double bond of the conjugated ring system of protochlorophyllide which is catalyzed by the multi-subunit enzyme dark operative protochlorophyllide oxidoreductase (DPOR). Subsequently, biosynthesis of all bacteriochlorophylls requires the reduction of the C7-C8 double bond by a second nitrogenase-like enzyme termed chlorophyllide oxidoreductase (COR). Mechanistically, DPOR and COR make use of a reductase component which links ATP hydrolysis to conformational changes. This dynamic switch protein is triggering the transient association between the reductase and the core catalytic protein complex, thereby facilitating the transduction of electrons via two [4Fe4S] clusters. X-ray crystallographic structural investigations in combination with biochemical experiments revealed the molecular basis of the underlying energy transduction mechanism. The unique nickel-containing tetrapyrrole cofactor F_{430} is located in the active site of methyl-coenzyme M reductase, which is catalyzing the final step of methane formation in methanogenic archaea. The nitrogenase-like protein NflH/NflD has been proposed to catalyze one or more ring reduction steps during the biosynthesis of F_{430}. The present working hypothesis mirrors a DPOR and COR related enzyme mechanism of NflH/NflD. Furthermore, *nfl*-encoded

G. Layer
Institut für Biochemie, Universität Leipzig, Brüderstraße 34, D-04103 Leipzig, Germany

J. Krausze
Institut für Pflanzenbiologie, Technische Universität Braunschweig, Spielmannstraße 7, D-38106 Braunschweig, Germany

J. Moser (✉)
Institut für Mikrobiologie, Technische Universität Braunschweig, Spielmannstraße 7, D-38106 Braunschweig, Germany
e-mail: j.moser@tu-bs.de

proteins were suggested as "simplified" ancestors lying basal in the phylogenetic tree between nitrogenase and DPOR/COR.

Abbreviations

DPOR	dark operative protochlorophyllide oxidoreductase
COR	chlorophyllide oxidoreductase
$(NifH)_2$	iron protein
Molybdenum-iron protein	MoFe protein or $(NifD/NifK)_2$
MCR	methyl-coenzyme M reductase

1   Introduction

The complex metalloenzyme nitrogenase is responsible for biological dinitrogen fixation, a process playing a central role in the global biogeochemical nitrogen cycle. This enzymatic conversion requires considerable energy input for the addition of electrons and protons to yield two molecules of ammonia (Fig. 1a). For more than 50 years, intense studies have been focusing on the understanding of the underlying enzyme mechanism, which allows for the sophisticated cleavage of the dinitrogen triple bond under ambient conditions. This is of special relevance, since the fueling of the related industrial dinitrogen fixing process (Haber-Bosch process at ~450 °C and >200 atm pressure) accounts for the consumption of approximately 1.4 % of the global energy demand (Lancaster et al. 2011).

The extensively characterized nitrogenase from *Azotobacter vinelandii* is composed of two subcomplexes, the iron protein $(NifH)_2$ and the molybdenum-iron protein (MoFe protein).

Homodimeric $(NifH)_2$ is bridged by a single intersubunit [4Fe4S] cluster and contains one ATP-binding site per polypeptide. The heterotetrameric MoFe protein $(NifD/NifK)_2$ is carrying two unique metal clusters per $(NifD/NifK)$-dimer: One [8Fe7S] cluster (termed P-cluster) located at the NifD/NifK-subunit interface and the [Mo7Fe9SC-homocitrate] cluster (termed FeMoco or M-cluster) which is buried within subunit NifD (Einsle et al. 2002; Kim and Rees 1992) (Fig. 1a). For catalytic dinitrogen reduction, the iron protein transiently forms a complex with the MoFe protein. This octameric complex $(NifH)_2(NifD/NifK)_2(NifH)_2$ then allows for the transfer of electrons from the [4Fe4S] cluster of $(NifH)_2$ to the P-cluster of $(NifD/NifK)_2$ in a strictly ATP-dependent process. Electrons are then further translocated and eventually accumulated on the FeMoco which is described as the substrate reduction site (Hoffman et al. 2014). These electron transfer processes are schematically depicted in Fig. 1a. Significant progress in understanding biological dinitrogen fixation was made in the last years: The interstitial carbide "holding together" the FeMoco was discovered, which is thought to be relevant for the developing of improved synthetic dinitrogen reducing catalysts (Magistrato et al. 2007; Kirchner et al. 2007; Bjornsson et al. 2015; Wiig et al. 2013). Furthermore, nitrogenase derived hydrocarbon formation from CO in analogy to the technical Fischer Tropsch synthesis was discovered and proposed as a potential strategy for fuel production in the future (Yang et al. 2011; Lee et al. 2012, 2015; Hu et al. 2011).

Photosynthesis performed by plants and bacteria is also a crucial biological process that
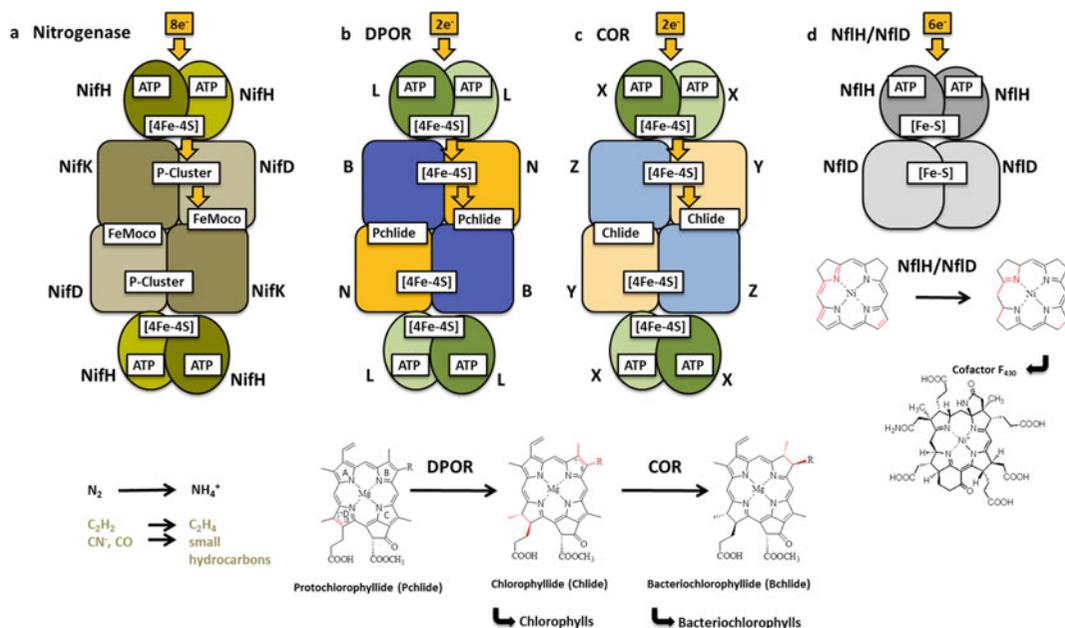
Fig. 1 Schematic comparison of transient octameric complexes involved in nitrogenase, DPOR and COR catalysis, hypothetical structural layout and catalysis of NflH/NflD. (**a–c**) Reductases (NifH)$_2$, L$_2$ and X$_2$ carrying an intersubunit [4Fe4S] cluster are shaded green. Core catalytic complexes (NifD/NifK)$_2$ of nitrogenase, (NB)$_2$ of nitrogenase-like DPOR and (YZ)$_2$ of nitrogenase-like COR are colored light brown/brown, yellow/blue and light yellow/light blue, respectively. ATP-driven electron transfer processes of nitrogenase (via [4Fe4S] cluster → P-cluster → FeMoco) and of DPOR or COR (via two [4Fe4S] clusters → Pchlide or Chlide) are schematically indicated. P-cluster [8Fe7S] cluster, FeMoco iron-molybdenum cofactor [Mo7Fe9SC-homocitrate] cluster also termed M-cluster. Reactions catalyzed by the respective systems are shown (*bottom*). (**d**) Hypothetical arrangement of subunits NflH and NflD, involved in cofactor F$_{430}$ biosynthesis. The NflH/NflD substrate is to be determined; therefore tetrapyrrole ring substituents have been omitted. R is either ethyl or vinyl

converts visible light into chemical energy to drive most ecosystems on earth (Hohmann-Marriott and Blankenship 2011). It is responsible for the fixation of $123 \cdot 10^9$ tons of carbon annually and the accumulation of vast quantities of organic deposits from which our current fossil fuels derive. Tetrapyrroles like chlorophylls and bacteriochlorophylls mainly contribute three distinct functions to photosynthesis: They are responsible for light capturing, the subsequent transfer of the obtained excitation energy, and they constitute the special pair involved in primary charge separation – a process which generates chemical energy in the form of a reduction potential (Muh et al. 2012). Primarily for the purpose of light capturing, huge quantities of (bacterio)chlorophylls are synthesized in photosynthetic organisms annually. Accordingly, these

pigments belong to the most abundant molecules on earth.

Chlorophyllide is a central hub metabolite for the biosynthesis of all chlorophylls and bacteriochlorophylls. For the synthesis of this molecule, the conjugated tetrapyrrole macrocycle of the pheoporphyrin protochlorophyllide must be stereospecifically reduced at the C17-C18 double bond on ring D (Fig. 1b). This chemically difficult two-electron trans-hydrogenation results in a significant change of the absorption characteristics. Protochlorophyllide reduction is either catalyzed by a single subunit enzyme that requires light for catalysis (light-dependent protochlorophyllide oxidoreductase; found in cyanobacteria, algae, gymnosperms and angiosperms) or alternatively by a dark operating system composed of three different polypeptides

termed dark operative protochlorophyllide oxidoreductase (DPOR; found in cyanobacteria, algae, gymnosperms and anoxic photosynthetic bacteria) (Fujita 1996; Reinbothe et al. 2010). DPOR performs an ATP-dependent catalysis and is composed of polypeptides L, N and B. The respective genes are denoted *chlL*, *chlN* and *chlB* in chlorophyll synthesizing organisms and *bchL*, *bchN* and *bchB* in bacteriochlorophyll synthesizing organisms. The metalloprotein DPOR is organized as a two-component system consisting of an ATP-dependent reductase L_2 and a catalytic subcomplex $(NB)_2$ both sharing a high degree of sequence and structural homology to the related nitrogenase proteins (Fig. 1b).

Bacteriochlorophylls characteristically differ from chlorophylls with respect to the redox state of the conjugated tetrapyrrole ring system (Burke et al. 1993) (Fig. 1c). Chlorophyllide oxidoreductase (COR) catalyzes the reaction step following DPOR catalysis in the bacteriochlorophyll biosynthesis pathway. COR enables for the chemically difficult reduction of the C7-C8 double bond of chlorophyllide. The resulting bacteriochlorophyllide molecule then is the central precursor for the biosynthesis of all bacteriochlorophylls (Nomata et al. 2006a). Mechanistically, DPOR and COR perform closely related two-electron reductions. However, accurate substrate recognition and adequate redox characteristics for the discrimination of the individual substrates are required. The ATP-dependent catalysis of COR requires polypeptides X, Y and Z, the respective genes have been denoted as *bchX*, *bchY* and *bchZ*. The subcomplex X_2 acts as an ATP-dependent reductase that interacts with the catalytic component $(YZ)_2$ of COR. The individual subunits of COR share an amino acid sequence identity of 31–35 % for subunit X and 15–22 % for subunits Y or Z when compared to the corresponding DPOR subunits (L, N or B). This homology might indicate a catalytic mechanism of COR which is closely related to DPOR catalysis.

Biological methane formation is a globally important process that is conducted exclusively by prokaryotes belonging to the domain of the archaea (Thauer et al. 2008). About one billion tons of methane are produced by methanogenic archaea annually. Almost half of the formed methane escapes into the atmosphere, where most of it is photochemically oxidized to CO_2 thus contributing to the greenhouse effect and global warming (Thauer 2011). Moreover, methane itself also acts as a very effective greenhouse gas. Methanogenic archaea also play a crucial role for the production of biogas with the formed methane representing the most important product.

During methane formation by methanogenic archaea the enzyme methyl-coenzyme M reductase (MCR) catalyzes the key step of the process, namely the reduction of the coenzyme M bound methyl group to gaseous methane (Ermler et al. 1997). In order to catalyze this reaction, MCR depends on the unique nickel-containing tetrapyrrole cofactor F_{430} as an essential prosthetic group. Cofactor F_{430} is an unusual tetrapyrrole with regard to the central metal ion (nickel), the presence of two additional rings attached to the central tetrapyrrole core (lactam ring E and cyclohexanone ring F) and the redox state of the ring system, which represents the most reduced tetrapyrrole found in nature (Fig. 1) (Friedmann et al. 1990). Although the biosynthesis route of cofactor F_{430} is currently largely unknown, it is obvious that an appropriate reductase is required for the reduction of the tetrapyrrole macrocycle (Pfaltz et al. 1987). Accordingly, it was proposed that unusual nitrogenase-like catalysis could be involved in the chemically difficult reduction of either two or three double bonds depending on the redox state of the actual substrate (Raymond et al. 2004). Assuming that the initial redox state before the reductase step corresponds to that of an isobacteriochlorin, the reductase has to catalyze the stereospecific two-electron reductions of the C12-C13 and the C18-C19 double bonds as well as the two electron-reduction of the C20-C4 diene to a single double bond as shown in Fig. 1d. Interestingly, all methanogenic archaea contain genes encoding a nitrogenase-like reductase system of unknown function consisting of a

reductase component NflH and a catalytic unit NflD (Nfl = <u>Nif</u>-like). Whereas NflH shows ~23 % amino acid sequence identity to NifH, NflD exhibits only ~6 % and ~5 % amino acid sequence identity to NifD and NifK, respectively. It was proposed that these enzymes might be responsible for the ring reduction steps during cofactor F_{430} biosynthesis (Bröcker et al. 2010a; Hu and Ribbe 2015).

2 Dark Operative Protochlorophyllide Oxidoreductase (DPOR)

2.1 Overview

DPOR is a highly oxygen sensitive two-component metalloprotein which performs the chemically difficult reduction of the conjugated ring system of protochlorophyllide in a strictly ATP-dependent reaction. Interdisciplinary approaches combining genetic, biochemical, spectroscopic, chemical and x-ray crystallographic methods contributed to the detailed understanding of the catalytic mechanism of DPOR. The enzyme is composed of two distinct subcomplexes: The homodimeric reductase L_2 is responsible for the initial uptake of an electron from a plant-type ferredoxin and for the subsequent transfer of this electron onto the heterotetrameric $(NB)_2$ subcomplex, where protochlorophyllide reduction takes place (Bröcker et al. 2008a). According to this oligomeric architecture, $(NB)_2$ contains two symmetry-related active site cavities for the binding of protochlorophyllide substrate molecules. At the same time, each NB half-tetramer has the ability to interact with one L_2 dimer, which gives rise to a theoretical $L_2(NB)_2L_2$ heterooctamer (Fig. 1b). Overall, the DPOR catalyzed two-electron reduction can be described as a reductive protonation of the C17-C18 double bond of protochlorophyllide which results in stereospecific chlorophyllide formation.

2.2 L_2 Is an ATP-Dependent Reductase

By analogy to $(NifH)_2$ of nitrogenase, L_2 forms a homodimer (M_r ~ 60 kDa) that is bridged via a symmetric [4Fe4S] cluster. This redox active metallocenter is coordinated by two conserved cysteine residues from each subunit (Bröcker et al. 2008a). The solved three-dimensional structure of the L_2 protein from *Rhodobacter sphaeroides* revealed this symmetrical dimer with one Mg^{2+}-coordinated ADP molecule bound to each subunit (Sarma et al. 2008). On the basis of mutational experiments, functional relevance of the highly conserved ATP cofactor binding motif called P-loop sequence (YGKGGIGK) and of the so called switch II region (LGDVVCGGF) was demonstrated (Bröcker et al. 2008a). In $(NifH)_2$ of nitrogenase, this switch II sequence acts as a conformational relay that communicates the binding of ATP to the [4Fe4S] cluster. Based on biochemical investigations and the overall structural similarity of L_2 and $(NifH)_2$ parallelism for the initial electron transfer steps of DPOR and nitrogenase was concluded.

2.3 The Catalytic $(NB)_2$ Complex: Substrate Recognition and Electron Transfer

The three-dimensional structures of substrate-free $(NB)_2$ (from the cyanobacterium *Thermosynechococcus elongatus*) (Bröcker et al. 2010a) and of $(NB)_2$ in the presence of protochlorophyllide (from the proteobacterium *Rhodobacter capsulatus*) (Muraki et al. 2010) revealed a high degree of structural similarity with respect to the quaternary and ternary structure. In both cases, $(NB)_2$ forms a symmetric heterotetramer which is comprised of two NB half tetramers. Sequences of N and B are paralogous and share an amino acid sequence identity of ~14 %, respectively. The overall fold of these individual subunits is rather similar which might indicate that sequences of N and B have evolved from a

common ancestor. The catalytic (NB)$_2$ complex binds two symmetry related [4Fe4S] centers at the NB interface, respectively. These clusters are asymmetrically coordinated by three cysteinyl ligands from N and one unusual aspartate ligand from B (Bröcker et al. 2010a; Muraki et al. 2010). When this aspartate was replaced by cysteine, the resulting mutant protein still assembled the [4Fe4S] cluster but almost no residual enzymatic activity was determined (Muraki et al. 2010). It was hypothesized, that the original aspartate ligand alters the redox potential of the [4Fe4S] cluster to a value below that of standard four cysteine ligated clusters. Obviously, the aspartate ligand is a prerequisite for protochlorophyllide reduction (Takano et al. 2011).

Each NB half tetramer contains one deeply buried protochlorophyllide binding site which is mainly located within subunit N. A partial unwinding of one helical segment from subunit B is required for the channeling of the substrate into the mainly hydrophobic binding pocket (Muraki et al. 2010; Moser et al. 2013). Furthermore, the C-terminal domain of subunit B is responsible for the closing of the active site cavity thereby preventing the organism from protochlorophyllide-induced photodynamic damage. DPOR substrate recognition has also been investigated using artificial protochlorophyllide derivatives. The related enzyme kinetics indicated that minor modifications on rings A, B, C and E (e.g. reducing the size or polarity of the side chain) were tolerated by the enzyme. However, substrate variants with an increased volume of the ring substituents and also variants with modifications on ring D were not converted. Obviously, the catalytic target on ring D is coordinated with high specificity (Bröcker et al. 2008b).

Each [4Fe4S] cluster is located at a distance of ~11 Å from the protochlorophyllide ring system (edge-to-edge distance), thereby allowing for rapid electron transfer to the substrate (Muraki et al. 2010; Moser et al. 2013; Bröcker et al. 2008b). However, the C17-C18 double bond of the substrate faces away from the [4Fe4S] cluster, arguing for an electron transfer process involving the conjugated ring systems A or B of the substrate (compare Fig. 1b and the ternary DPOR structure depicted in Fig. 2a).

With respect to the regio- and stereo-specific reduction of the substrate, a direct protonation at C17 via a highly conserved aspartate residue was concluded. By contrast, the C18 protonation is mediated by 'substrate assisted catalysis'. A water molecule is well-positioned by combined interaction with a histidine residue and the C18 propionate of protochlorophyllide. This water then facilitates the trans-specific C18 protonation (Moser et al. 2013).

2.4 Trapping the Ternary Complex Reveals the Catalytic Redox Cycle of DPOR

As in nitrogenase, it was demonstrated that subcomplexes L$_2$ and (NB)$_2$ solely perform a transient protein protein interaction during the time course of DPOR catalysis. The dynamic interplay of these two components is triggered by ATP hydrolysis in a process that facilitates the appropriate timing of the electron transfer step between L$_2$ and (NB)$_2$. The trapping of the ternary transition state complex allowed for the detailed understanding of DPOR subcomplex interaction and the involved protein dynamics. Biochemical experiments in the presence of the non-hydrolysable ATP analog MgADP-AlF$_3$ efficiently produced the octameric L$_2$(NB)$_2$L$_2$ protein. This stabilized complex is impaired in electron transfer to (NB)$_2$ which confirms the coupling between ATP hydrolysis and electron transfer. It was studied in great detail and subsequently the crystal structure of this ~360 kDa complex was resolved (Moser et al. 2013; Bröcker et al. 2010b) (Fig. 2a, c).

The catalytic cycle of DPOR involves a large number of individual steps. Biochemical, EPR spectroscopic methods and the comparative analysis of the available X-ray crystallographic structures provided a detailed picture of the order of events. The catalytic cycle of DPOR is summarized in Fig. 3, individual states that have been characterized by EPR spectroscopy are
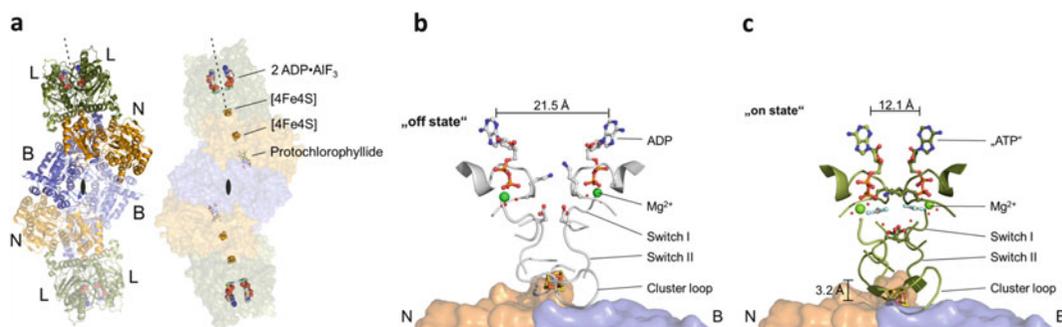
Fig. 2 Three-dimensional structure of the ternary DPOR complex and dynamic switch mechanism of L_2. (**a**) L_2 carrying an intersubunit [4Fe4S] cluster is a dynamic switch protein triggering the transient interaction with (NB)$_2$. This results in the transfer of a single electron onto the substrate protochlorophyllide via a second [4Fe4S] cluster located on (NB)$_2$ (PDB ID code 2YNM). (**b**, **c**) Nucleotide-dependent switch mechanism of DPOR subcomplex L_2, minor residues are omitted for clarity. (**b**) Key secondary structural elements of L_2 in the "off state" conformation (*grey*), observed in the presence of ADP (PDB ID code 3FWY). (**c**) Identical structural elements of L_2 in the "on state" conformation (*green*), observed in the presence of the ATP analog ADP-AlF$_3$. Bound nucleotide cofactors (with indicated distances), switch region I, switch region II and the cluster loop responsible for the dynamic repositioning of the [4Fe4S] cluster of L_2 are indicated. Movement of the [4Fe4S] cluster of L_2 over a distance of 3.2 Å is indicated. Panel a modified according to (Moser et al. 2013)

highlighted grey (Bröcker et al. 2010b). The catalytic (NB)$_2$ complex was shown to have a high affinity for its protochlorophyllide substrate (Bröcker et al. 2008b; Nomata et al. 2008). Obviously, the initial binding of the substrate might be a critical step in vivo to overcome the well described phototoxicity of the DPOR substrate (Walther et al. 2009). In the laboratory, DPOR experiments are performed in the presence of the artificial electron donor dithionite which results in the reduction of the [4Fe4S] cluster of L_2. This single electron reduction does not require the presence of (NB)$_2$ and/or ATP. It was shown that the reduced and ATP-charged L_2 protein is the only electron donor with the ability to transfer electrons onto (NB)$_2$ during DPOR catalysis (Bröcker et al. 2008a; Nomata et al. 2008).

In vivo, a ferredoxin acts as the natural electron donor and transfers one electron onto the [4Fe4S] cluster of L_2 (**I**) (Bröcker et al. 2008a; Nomata et al. 2006b). L_2 is a dynamic switch protein that links the hydrolysis of ATP to significant conformational rearrangements of the overall ternary protein structure. The L_2 protein can be characterized by two different states: the "off state" in the presence of ADP which is not able to form a complex with (NB)$_2$ and the "on

state" which possesses a high affinity for (NB)$_2$ (Bröcker et al. 2008a; Sarma et al. 2008; Nomata et al. 2006b). Accordingly, binding of two ATP molecules to L_2 (**II**) results in conformational alterations (compare protein dynamics depicted in Fig. 2b, c) and the subsequent formation of the ternary DPOR complex (**III**) in the presence of (NB)$_2$ containing two protochlorophyllide molecules (Bröcker et al. 2010b; Watzlich et al. 2009). ATP hydrolysis facilitates the translocation of one electron from the [4Fe4S] cluster of L_2 onto the low redox potential [4Fe4S] cluster located on (NB)$_2$ (**IV**). The reduced (NB)$_2$ protein then has the ability to transfer a single electron onto the substrate. This process is induced by the dynamic repositioning of the [4Fe4S] cluster of L_2 by 3.2 Å as a result of ternary complex formation (compare the "off state" and the "on state" conformation depicted in Fig. 2b, c) (Moser et al. 2013). In the L_2(NB)$_2$L$_2$ complex rapid electron transfer is facilitated due to an edge-to-edge distance of 14.1 Å between the [4Fe4S] cluster of L_2 and (NB)$_2$. Presence of ADP then triggers the dissociation of the L_2(NB)$_2$L$_2$ complex (**V**). Overall, two consecutive rounds of this redox catalytic cycle are required so that the dynamic switch protein L_2
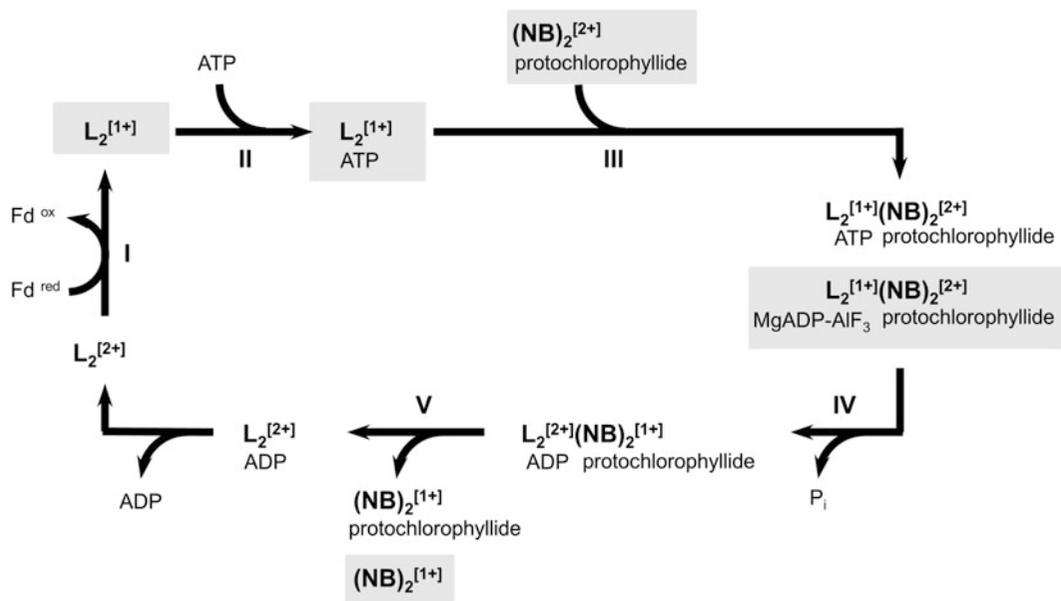
Fig. 3 Redox cycle of DPOR catalysis. Schematic model for the electron transfer processes and dynamic subunit interaction during ATP-driven DPOR catalysis. Five intermediates were confirmed by EPR spectroscopy (*highlighted grey*), the individual redox state is indicated [1+] for reduced and [2+] for oxidized [4Fe4S] clusters. According to this redox cycle two consecutive single electron reductions of $(NB)_2$ are required to provide the two electrons necessary for protochlorophyllide reduction

can supply the two electrons which are required for the reduction of the substrate. These individual electron transfer steps go along with the stereospecific addition of two protons to C17 and C18 of protochlorophyllide in the active site of DPOR (Muraki et al. 2010; Moser et al. 2013). Only recently, formation of the proposed substrate radical after single electron transfer was confirmed by EPR spectroscopy (Nomata et al. 2014).

The spectrum of artificial 'small-molecule substrates' of DPOR was demonstrated in correlation with those of the related nitrogenase system (Fig. 1a). DPOR does not catalyze the conversion of 'complex' nitrogenase substrates like N_2 or CO since these reductions require more than two electrons. By contrast, DPOR enables for the two-electron reduction of N_3^- or N_2H_4 to NH_3 analogously as also determined for nitrogenase (Moser et al. 2013). These results might indicate that DPOR and nitrogenase make use of a closely related energy transduction mechanism. According to this, DPOR might be used as an important tool to gain further insight into the mechanism of dinitrogen fixation.

3 Chlorophyllide Oxidoreductase (COR)

3.1 Overview

Synthesis of all bacteriochlorophylls involves the ATP-driven reduction of the C7-C8 double bond of chlorophyllide by the multi-subunit enzyme COR (Fig. 1c). Structural information for this oxygen sensitive enzyme system is not available to date. However, combined mutational and spectroscopic analyses, the homology to the DPOR system and also chimeric COR/DPOR experiments revealed a detailed picture of COR catalysis. The reductive protonation of chlorophyllide is based on the transient interaction of subcomplex X_2 with the catalytic $(YZ)_2$ protein for the electron transfer via two redox active iron-sulfur centers.

Only recently, it was demonstrated that COR systems from *R. capsulatus*, *R. sphaeroides*, *Rhodopseudomonas palustris*, *Chlorobaculum tepidum* and *Roseiflexus castenholzii* are also capable of the reduction of the $C8^1$-$C8^2$ double bond (8-vinyl group) of chlorophyllide (Tsukatani et al. 2013a; Harada et al. 2014). Besides these bifunctional enzymes, it was also demonstrated that the COR enzyme from *Heliobacterium modesticaldum* is able to catalyze the direct 8-ethylidene group formation during the biosynthesis of bacteriochlorophyll *g* (Tsukatani et al. 2013b). In Fig. 4 these additional COR activities, the reductive 1,2 protonation at $C8^1$-$C8^2$ and the reductive 1,4 protonation at $C7$-$C8^2$ of chlorophyllide are shown. Mechanistically, these 'additional' enzymatic activities of specific COR enzymes might be closely related to the enzyme catalysis of NflH/NflD described in chapter 4.

3.2   The Reductase X_2 Resembles L_2

Size exclusion experiments for the COR reductase protein from the purple bacterium *Roseobacter denitrificans* indicated a homodimeric quaternary structure (Kiesel et al. 2015). Sequence comparisons for X proteins with the related L sequences of DPOR revealed that both cysteinyl ligands responsible for the formation of the redox-active intersubunit [4Fe4S] cluster of L_2 are fully conserved among all X proteins (Burke et al. 1993). EPR measurements for X_2 displayed the characteristic signal of a [4Fe4S] cluster (Kiesel et al. 2015; Kim et al. 2008). Besides this, the key amino acid residues for the dynamic switch mechanism of L_2 (responsible for ATP binding and hydrolysis, signal transduction) are highly conserved in all sequences of X proteins (Watzlich et al. 2009). According to these findings, parallelism for the electron transfer steps of X_2 and L_2 via an intersubunit [4Fe4S] cluster was concluded. To some extent, this hypothesis was experimentally confirmed. The reductase L_2 of DPOR was substituted with subunit X_2 from COR in a DPOR activity assay in the presence of protochlorophyllide. The resulting chimeric enzyme (e.g. X_2 from *R. denitrificans* / $(NB)_2$ from *Chlorobaculum tepidum*) revealed substantial protochlorophyllide reducing activity (Watzlich et al. 2009). These experiments might indicate that the ATP-driven electron transfer mechanism of X_2 mirrors the well described dynamic switch mechanism of L_2. Furthermore, it was speculated that the docking faces of X_2 and $(YZ)_2$ and the related protein protein interplay of COR might resemble that of the DPOR system.

3.3   The Catalytic $(YZ)_2$ Complex: Substrate Recognition and Electron Transfer

Size exclusion experiments resulted in a native molecular mass of ~ 280 kDa, indicative of a heterotetrameric $(YZ)_2$ complex in analogy to the related DPOR system. Substrate binding experiments for the *R. denitrificans* $(YZ)_2$ protein revealed the tight binding of the photolabile chlorophyllide molecule and suggested the presence of two substrate binding sites (Kiesel et al. 2015). COR activity experiments using a series of chemically modified substrates tentatively revealed the binding of chlorophyllide in a buried active site cavity (Kiesel et al. 2015). The COR enzyme efficiently reduced the C7-C8 double bond of substrates containing modified substituents on ring systems A, C, and E. By contrast, artificial substrates modified at the distantly located propionate side chain of ring D were not a substrate of COR. According to this, a closely related substrate binding mode for chlorophyllide and protochlorophyllide with regard to the ring orientation in the active site of COR and DPOR was hypothesized. It was concluded that both systems make use of an evolutionary conserved electron insertion path via ring systems A or B. Accordingly, the active site of COR would make use of two alternative proton donors (when compared to DPOR) for the regio- and stereospecific reduction of ring B. With respect to the COR-dependent 8-ethylidene group formation in *H. modesticaldum* (Fig. 4), this alternative
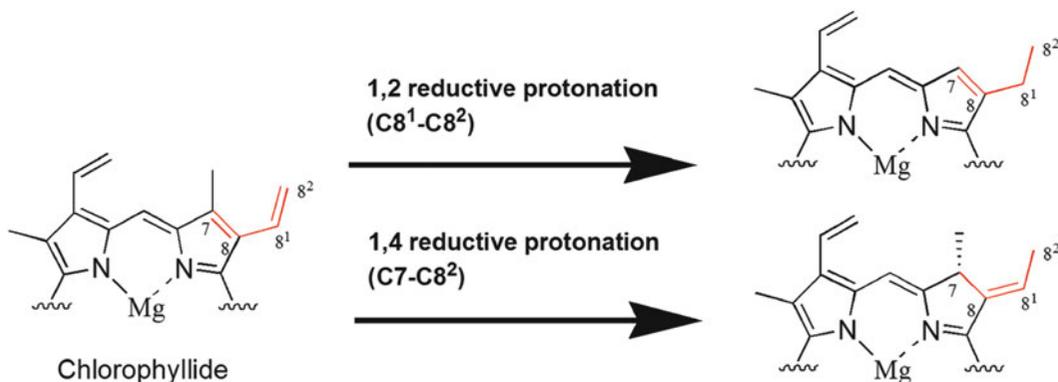
Fig. 4 Additional or alternative enzymatic activities of COR enzymes involved in bacteriochlorophyll *a* or bacteriochlorophyll *g* biosynthesis. COR was also proposed as an alternative 8-vinyl reductase involved in bacteriochlorophyll *a* biosynthesis in *R. capsulatus*, *R. sphaeroides*, *R. palustris*, *C. tepidum* and *R. castenholzii* (*top*). Furthermore, COR catalyzed 8-ethylidene group formation during bacteriochlorophyll *g* biosynthesis was demonstrated for *H. modesticaldum* (*bottom*). These two-electron reduction processes can be described as an 1,2 reductive protonation or as an 1,4 reductive protonation

COR activity is based on the 1,4-hydrogenation of a diene system at C7 and $C8^2$ of chlorophyllide (instead of the well described 1,2-hydrogenation of C7-C8). This catalysis is based on an orthologous COR enzyme and it might only require the spatial rearrangement of a single proton donor ($C8^2$ instead of C8). With respect to the bifunctional COR enzymes, performing the hydrogenation of C7-C8 in parallel with the hydrogenation of the $C8^1$-$C8^2$ double bond, an overall of four specific proton donors would be required for the catalyzed four electron reduction (Kiesel et al. 2015).

EPR experiments for $(YZ)_2$ revealed a characteristic signal for a [4Fe4S] cluster. However, mutational experiments were indicative of four cysteinyl cluster ligands. This was in clear contrast to the related DPOR system where the [4Fe4S] cluster of $(NB)_2$ makes use of three cysteine and one unusual aspartate ligand, which might be relevant for the "tuning" of the respective redox potential (Kondo et al. 2011). For the $(YZ)_2$ protein from *R. denitrificans* a DPOR-like three cysteine/one aspartate ligation pattern for the catalytic [4Fe4S] cluster was implemented by mutagenesis. This mutant protein revealed artificial [4Fe4S] cluster formation as indicated by EPR experiments. However, no enzymatic COR activity was determined in the presence of this cluster variant which was ascribed to an inappropriate redox potential of the assembled [4Fe4S] cluster (Kiesel et al. 2015).

3.4 The Ternary COR Complex Reveals Parallelism for the COR and DPOR Redox Cycle

The ATP-dependent conversion of chlorophyllide strictly requires the presence of X_2 and $(YZ)_2$. In vitro experiments using the artificial reducing agent dithionite clearly indicated that the electron transferring reductase L_2 of DPOR cannot substitute for the COR specific reductase X_2. Biophysical and biochemical investigations suggested [4Fe-4S] cluster-dependent redox catalysis by analogy to the related DPOR system. Accordingly, ternary COR complex formation via the nucleotide-dependent reductase X_2 was investigated in the presence of the substrate and the ATP analog MgADP-AlF_4^-. MgADP in combination with the inorganic compound AlF_4^- is mimicking the ATP molecule in the transition state of nucleotide

hydrolysis. From the stoichiometry of the assembled COR complex an octameric $X_2(YZ)_2X_2$ complex was deduced (Fig. 1c) (Kiesel et al. 2015). A catalytic redox cycle for the COR system was proposed by analogy to the mechanism depicted in Fig. 3.

4 NflH/NflD

4.1 Overview

As mentioned in the introduction, another nitrogenase-like enzyme system besides DPOR and COR was identified on the basis of bioinformatic analysis of prokaryotic genomes and was termed Nif-like (Nfl) due to its homology to nitrogenase (Raymond et al. 2004). The Nfl-system also consists of two components: NflH which was proposed as a homolog of the reductases NifH, L and X and the proposed catalytic component NflD which is related to the paralogous protein sequences NifD/NifK, N/B and Y/Z, respectively. Based on phylogenetic analyses it was proposed that the Nfl-system might represent an ancestor of nitrogenase and DPOR/COR. These enzymes could have evolved from this ancestor through paralogous gene duplication and subsequent divergence (Boyd and Peters 2013). Interestingly, the Nfl proteins occur in all methanogenic archaea. In 2007, the expression pattern of *nflH* and *nflD* in *Methanocaldococcus jannaschii* was investigated. It was found that these genes are expressed constitutively and, therefore, it was concluded that NflH and NflD fulfill some housekeeping function essential for methanogens such as the biosynthesis of cofactor F_{430}. Further, it was shown by co-purification and bacterial two-hybrid studies that NflH and NflD interact with each other *in vivo* (Staples et al. 2007). However, the precise function of NflH/NflD remains to be elucidated. Therefore, the proposed function and mechanism of NflH/NflD described below are based on the homology to nitrogenase, DPOR and COR.

4.2 The Proposed Reductase NflH

Based on amino acid sequence similarity with NifH, L and X proteins it is very likely that NflH also adopts a homodimeric structure carrying a bridging intersubunit [4Fe4S] cluster. The amino acid sequences of NflH proteins contain three highly conserved cysteine residues with two of them corresponding to the iron-sulfur cluster ligands of NifH, L and X. Moreover, NflH sequences also exhibit the characteristic P-loop sequence (YGKGGIGK) and the so called switch II region. However, the switch II region of several NflH is slightly altered. Instead of the highly conserved motif LGDVVCGGF found in NifH, L and X, the amino acid sequences of some NflH proteins carry a PGDIVCGGF motif. The leucine residue within the former motif was shown to be essential for the dynamic switch mechanism (Bröcker et al. 2010b; Lanzilotta et al. 1996). Since this residue is replaced by a proline in these NflH sequences, the mechanism of ATP-dependent conformational changes required for electron transfer from the reductase to the catalytic component might be different in some of the Nfl systems. Nevertheless, it is obvious that NflH and NflD have to interact with each other for electron transfer, and such an interaction was indeed observed to occur in vivo (Staples et al. 2007).

4.3 The Proposed Catalytic (NflD)$_2$ Complex

The amino acid sequences of NflD proteins exhibit similarity to the sequences of NifD/NifK, N/B and Y/Z. Thus, it is assumed that NflD represents the catalytic component of the Nfl system. However, in contrast to nitrogenase, DPOR and COR, the catalytic component of the Nfl system is not a heterotetramer, since there is no second subunit besides NflD, but rather a homodimer or alternatively a homotetramer. Within the amino acid sequences of NflD there are two highly conserved cysteine residues which correspond to two of the iron-sulfur cluster
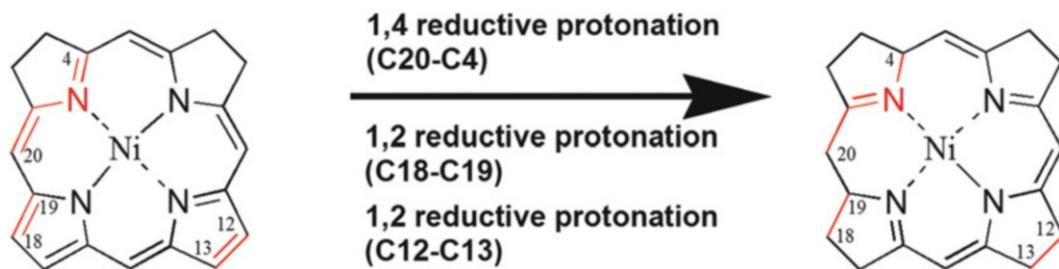
Fig. 5 Proposed NflH/NflD catalysis. The ring reduction steps include the 1,4 reductive protonation of a diene system at C20-C4 and the 1,2 reductive protonation at C18-C19 and at C12-C13. The precise NflH/NflD substrate is to be determined; accordingly the individual tetrapyrrole ring substituents have been omitted

ligating cysteines of NifD, N and Y. Thus, it is possible that the proposed NflD dimer contains a bridging intersubunit [4Fe4S] cluster. Considering a homodimeric $(NflD)_2$ structure, an intriguing question will be whether the asymmetric substrate binds to an active site located between the two monomers or whether there are two identical active sites located within each monomer. In any case, probably several different proton donating residues are required for the three reductive protonation steps during cofactor F_{430} biosynthesis.

4.4 Protein Protein Interactions and Proposed Reaction Catalyzed by NflH/NflD

Based on the assumptions described above, it is reasonable to suggest that the homodimeric reductase $(NflH)_2$ interacts with the homodimeric catalytic $(NflD)_2$ component to form a $(NflH)_2/(NflD)_2$ complex (Fig. 1d). Such a complex would represent a minimal version in order to achieve nitrogenase-like catalysis (Hu and Ribbe 2015). The necessity of ATP-binding and hydrolysis for complex formation and electron transfer from $(NflH)_2$ to $(NflD)_2$ is very likely but remains to be shown experimentally.

The reaction catalyzed within the catalytic $(NflD)_2$ is most likely comprised of two 1,2 reductive protonation steps, in order to stereospecifically reduce the C12-C13 and C18-C19 double bonds, and one 1,4 reductive protonation in

order to convert the C20-C4 diene to a single C1-N1 double bond (Fig. 5). Structural and biochemical analysis of the NflH/NflD system in the future will reveal the amino acid residues involved in these reactions and will possibly provide insights into the evolutionary relationship between NflH/NflD, DPOR, COR and nitrogenase.

5 Concluding Remarks

The enzymatic mechanisms of DPOR and COR (and most likely of NflH/NflD) indicate a high degree of plasticity for the reductive protonation of differing tetrapyrrole molecules. However, the individual substrates differ with respect to the redox state of the conjugated ring system. A deeply buried substrate binding pocket devoid of water molecules or unspecific proton donors might be nature's design to avoid unspecific 'tetrapyrrole over-reduction'. The overall amount of precisely positioned proton donors might account for the individual two, four or six electron reductions catalyzed by DPOR/COR, COR or NflH/NflD, respectively. Furthermore, the redox potential of the involved [4Fe4S] clusters (reductase and catalytic component) might be also relevant for the respective tetrapyrrole reducing activity. With respect to this, the further elucidation of the catalytic differences of DPOR, COR and NflH/NflD also might have implications for the further understanding of nitrogenase catalysis.

References

Bjornsson R, Neese F, Schrock RR, Einsle O, DeBeer S (2015) The discovery of Mo(III) in FeMoco: reuniting enzyme and model chemistry. J Biol Inorg Chem 20(2):447–460. doi:10.1007/s00775-014-1230-6

Boyd ES, Peters JW (2013) New insights into the evolutionary history of biological nitrogen fixation. Front Microbiol 4:201. doi:10.3389/fmicb.2013.00201

Bröcker MJ, Virus S, Ganskow S, Heathcote P, Heinz DW, Schubert WD, Jahn D, Moser J (2008a) ATP-driven reduction by dark-operative protochlorophyllide oxidoreductase from Chlorobium tepidum mechanistically resembles nitrogenase catalysis. J Biol Chem 283(16):10559–10567. doi:10.1074/jbc.M708010200

Bröcker MJ, Watzlich D, Uliczka F, Virus S, Saggu M, Lendzian F, Scheer H, Rudiger W, Moser J, Jahn D (2008b) Substrate recognition of nitrogenase-like dark operative protochlorophyllide oxidoreductase from Prochlorococcus marinus. J Biol Chem 283 (44):29873–29881

Bröcker MJ, Schomburg S, Heinz DW, Jahn D, Schubert WD, Moser J (2010a) Crystal structure of the nitrogenase-like dark operative protochlorophyllide oxidoreductase catalytic complex (ChlN/ChlB)₂. J Biol Chem 285(35):27336–27345. doi:10.1074/jbc.M110.126698

Bröcker MJ, Wätzlich D, Saggu M, Lendzian F, Moser J, Jahn D (2010b) Biosynthesis of (bacterio) chlorophylls: ATP-dependent transient subunit interaction and electron transfer of dark operative protochlorophyllide oxidoreductase. J Biol Chem 285(11):8268–8277

Burke DH, Hearst JE, Sidow A (1993) Early evolution of photosynthesis: clues from nitrogenase and chlorophyll iron proteins. Proc Natl Acad Sci U S A 90 (15):7134–7138

Einsle O, Tezcan FA, Andrade SL, Schmid B, Yoshida M, Howard JB, Rees DC (2002) Nitrogenase MoFe-protein at 1.16 Å resolution: a central ligand in the FeMo-cofactor. Science 297(5587):1696–1700. doi:10.1126/science.1073877

Ermler U, Grabarse W, Shima S, Goubeaud M, Thauer RK (1997) Crystal structure of methyl-coenzyme M reductase: the key enzyme of biological methane formation. Science 278(5342):1457–1462

Friedmann HC, Klein A, Thauer RK (1990) Structure and function of the nickel porphinoid, coenzyme F430 and of its enzyme, methyl coenzyme M reductase. FEMS Microbiol Rev 7(3-4):339–348

Fujita Y (1996) Protochlorophyllide reduction: a key step in the greening of plants. Plant Cell Physiol 37(4):411–421

Harada J, Mizoguchi T, Tsukatani Y, Yokono M, Tanaka A, Tamiaki H (2014) Chlorophyllide a oxidoreductase works as one of the divinyl reductases specifically involved in bacteriochlorophyll a biosynthesis. J Biol Chem 289(18):12716–12726. doi:10.1074/jbc.M113.546739

Hoffman BM, Lukoyanov D, Yang ZY, Dean DR, Seefeldt LC (2014) Mechanism of nitrogen fixation by nitrogenase: the next stage. Chem Rev 114(8):4041–4062. doi:10.1021/cr400641x

Hohmann-Marriott MF, Blankenship RE (2011) Evolution of photosynthesis. Annu Rev Plant Biol 62:515–548. doi:10.1146/annurev-arplant-042110-103811

Hu Y, Ribbe MW (2015) Nitrogenase and homologs. J Biol Inorg Chem 20(2):435–445. doi:10.1007/s00775-014-1225-3

Hu Y, Lee CC, Ribbe MW (2011) Extending the carbon chain: hydrocarbon formation catalyzed by vanadium/molybdenum nitrogenases. Science 333(6043):753–755. doi:10.1126/science.1206883

Kiesel S, Watzlich D, Lange C, Reijerse E, Brocker MJ, Rudiger W, Lubitz W, Scheer H, Moser J, Jahn D (2015) Iron-sulfur cluster-dependent catalysis of chlorophyllide a oxidoreductase from Roseobacter denitrificans. J Biol Chem 290(2):1141–1154. doi:10.1074/jbc.M114.617761

Kim J, Rees DC (1992) Crystallographic structure and functional implications of the nitrogenase molybdenum-iron protein from Azotobacter vinelandii. Nature 360(6404):553–560

Kim EJ, Kim JS, Lee IH, Rhee HJ, Lee JK (2008) Super-oxide generation by chlorophyllide a reductase of Rhodobacter sphaeroides. J Biol Chem 283 (7):3718–3730

Kirchner B, Wennmohs F, Ye S, Neese F (2007) Theoretical bioinorganic chemistry: the electronic structure makes a difference. Curr Opin Chem Biol 11(2):134–141. doi:10.1016/j.cbpa.2007.02.026

Kondo T, Nomata J, Fujita Y, Itoh S (2011) EPR study of 1Asp-3Cys ligated 4Fe-4S iron-sulfur cluster in NB-protein (BchN-BchB) of a dark-operative protochlorophyllide reductase complex. FEBS Lett 585(1):214–218. doi:10.1016/j.febslet.2010.11.044

Lancaster KM, Roemelt M, Ettenhuber P, Hu Y, Ribbe MW, Neese F, Bergmann U, DeBeer S (2011) X-ray emission spectroscopy evidences a central carbon in

the nitrogenase iron-molybdenum cofactor. Science 334(6058):974–977. doi:10.1126/science.1206445

Lanzilotta WN, Fisher K, Seefeldt LC (1996) Evidence for electron transfer from the nitrogenase iron protein to the molybdenum-iron protein without MgATP hydrolysis: characterization of a tight protein-protein complex. Biochemistry 35(22):7188–7196. doi:10.1021/bi9603985

Lee CC, Hu Y, Ribbe MW (2012) ATP-independent formation of hydrocarbons catalyzed by isolated nitrogenase cofactors. Angew Chem Int Ed Engl 51(8):1947–1949. doi:10.1002/anie.201108916

Lee CC, Hu Y, Ribbe MW (2015) Insights into hydrocarbon formation by nitrogenase cofactor homologs. MBio 6(2). doi:10.1128/mBio.00307-15

Magistrato A, Robertazzi A, Carloni P (2007) Nitrogen fixation by a molybdenum catalyst mimicking the function of the nitrogenase enzyme: a critical evaluation of DFT and solvent effects. J Chem Theory Comput 3(5):1708–1720. doi:10.1021/ct700094y

Moser J, Lange C, Krausze J, Rebelein J, Schubert WD, Ribbe MW, Heinz DW, Jahn D (2013) Structure of ADP-aluminium fluoride-stabilized protochlorophyllide oxidoreductase complex. Proc Natl Acad Sci U S A 110(6):2094–2098. doi:10.1073/pnas.1218303110

Muh F, Glockner C, Hellmich J, Zouni A (2012) Light-induced quinone reduction in photosystem II. Biochim Biophys Acta 1817(1):44–65. doi:10.1016/j.bbabio.2011.05.021

Muraki N, Nomata J, Ebata K, Mizoguchi T, Shiba T, Tamiaki H, Kurisu G, Fujita Y (2010) X-ray crystal structure of the light-independent protochlorophyllide reductase. Nature 465(7294):110–114 http://www.nature.com/nature/journal/v465/n7294/suppinfo/nature08950_S1.html

Nomata J, Mizoguchi T, Tamiaki H, Fujita Y (2006a) A second nitrogenase-like enzyme for bacteriochlorophyll biosynthesis: reconstitution of chlorophyllide a reductase with purified X-protein (BchX) and YZ-protein (BchY-BchZ) from Rhodobacter capsulatus. J Biol Chem 281(21):15021–15028. doi:10.1074/jbc.M601750200

Nomata J, Kitashima M, Inoue K, Fujita Y (2006b) Nitrogenase Fe protein-like Fe–S cluster is conserved in L-protein (BchL) of dark-operative protochlorophyllide reductase from Rhodobacter capsulatus. FEBS Lett 580(26):6151–6154 http://dx.doi.org/10.1016/j.febslet.2006.10.014

Nomata J, Ogawa T, Kitashima M, Inoue K, Fujita Y (2008) NB-protein (BchN–BchB) of dark-operative protochlorophyllide reductase is the catalytic component containing oxygen-tolerant Fe–S clusters. FEBS Lett 582(9):1346–1350. doi:10.1016/j.febslet.2008.03.018

Nomata J, Kondo T, Mizoguchi T, Tamiaki H, Itoh S, Fujita Y (2014) Dark-operative protochlorophyllide oxidoreductase generates substrate radicals by an iron-sulphur cluster in bacteriochlorophyll biosynthesis. Sci Rep 4:5455. doi:10.1038/srep05455 http://www.nature.com/articles/srep05455#supplementary-information

Pfaltz A, Kobelt A, Huster R, Thauer RK (1987) Biosynthesis of coenzyme F430 in methanogenic bacteria. Identification of 15,17(3)-seco-F430-17(3)-acid as an intermediate. Eur J Biochem 170(1-2):459–467

Raymond J, Siefert JL, Staples CR, Blankenship RE (2004) The natural history of nitrogen fixation. Mol Biol Evol 21(3):541–554. doi:10.1093/molbev/msh047

Reinbothe C, El Bakkouri M, Buhr F, Muraki N, Nomata J, Kurisu G, Fujita Y, Reinbothe S (2010) Chlorophyll biosynthesis: spotlight on protochlorophyllide reduction. Trends Plant Sci 15(11):614–624 doi:S1360-1385(10)00155-X [pii] 1016/j.tplants.2010.07.002

Sarma R, Barney BM, Hamilton TL, Jones A, Seefeldt LC, Peters JW (2008) Crystal structure of the L protein of Rhodobacter sphaeroides light-independent protochlorophyllide reductase with MgADP bound: a homologue of the Nitrogenase Fe Protein. Biochemistry 47(49):13004–13015. doi:10.1021/bi801058r

Staples CR, Lahiri S, Raymond J, Von Herbulis L, Mukhophadhyay B, Blankenship RE (2007) Expression and association of group IV nitrogenase NifD and NifH homologs in the non-nitrogen-fixing archaeon Methanocaldococcus jannaschii. J Bacteriol 189(20):7392–7398. doi:10.1128/JB.00876-07

Takano Y, Yonezawa Y, Fujita Y, Kurisu G, Nakamura H (2011) Electronic structures of a [4Fe–4S] cluster, [Fe4S4(SCH3)3(CH3COO)], in dark-operative protochlorophyllide oxidoreductase (DPOR). Chem Phys Lett 503(4–6):296–300 http://dx.doi.org/10.1016/j.cplett.2011.01.026

Thauer RK (2011) Anaerobic oxidation of methane with sulfate: on the reversibility of the reactions that are catalyzed by enzymes also involved in methanogenesis from CO_2. Curr Opin Microbiol 14(3):292–299. doi:10.1016/j.mib.2011.03.003

Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R (2008) Methanogenic archaea: ecologically relevant differences in energy conservation. Nat Rev Microbiol 6(8):579–591. doi:10.1038/nrmicro1931

Tsukatani Y, Yamamoto H, Harada J, Yoshitomi T, Nomata J, Kasahara M, Mizoguchi T, Fujita Y, Tamiaki H (2013a) An unexpectedly branched biosynthetic pathway for bacteriochlorophyll b capable of absorbing near-infrared light. Sci Rep 3:1217. doi:10.1038/srep01217

Tsukatani Y, Yamamoto H, Mizoguchi T, Fujita Y, Tamiaki H (2013b) Completion of biosynthetic pathways for bacteriochlorophyll g in Heliobacterium modesticaldum: the C8-ethylidene group formation. Biochim Biophys Acta 1827(10):1200–1204. doi:10.1016/j.bbabio.2013.06.007

Walther J, Brocker MJ, Watzlich D, Nimtz M, Rohde M, Jahn D, Moser J (2009) Protochlorophyllide: a new

photosensitizer for the photodynamic inactivation of Gram-positive and Gram-negative bacteria. FEMS Microbiol Lett 290(2):156–163

Watzlich D, Brocker MJ, Uliczka F, Ribbe M, Virus S, Jahn D, Moser J (2009) Chimeric nitrogenase-like enzymes of (bacterio)chlorophyll biosynthesis. J Biol Chem 284(23):15530–15540

Wiig JA, Lee CC, Hu Y, Ribbe MW (2013) Tracing the interstitial carbide of the nitrogenase cofactor during substrate turnover. J Am Chem Soc 135 (13):4982–4983. doi:10.1021/ja401698d

Yang ZY, Dean DR, Seefeldt LC (2011) Molybdenum nitrogenase catalyzes the reduction and coupling of CO to form hydrocarbons. J Biol Chem 286(22):19417–19421. doi:10.1074/jbc.M111.229344

Index