Alejandro G. Marangoni

Kinetic Analysis of Food Systems

Springer

Kinetic Analysis of Food Systems

Alejandro G. Marangoni

Kinetic Analysis of Food Systems

Springer

Alejandro G. Marangoni
Department of Food Science
University of Guelph
Guelph
Ontario
Canada

ISBN 978-3-319-51291-4 ISBN 978-3-319-51292-1 (eBook)
DOI 10.1007/978-3-319-51292-1

Library of Congress Control Number: 2017932900

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedicated to my colleagues who remind me that the scientific endeavor is a privilege and an honor.

Introduction: Trials and Tribulations in the Study of Complex, Multicomponent, Multiphasic, Nonequilibrium Systems

> The computer model is to the mind what the telescope and microscope are to the eye. (Heinz Pagels)

Food systems encompass anything from whole organisms and living tissue to nanostructured soft materials: potatoes to cheese. One characteristic all these systems share is complexity. All food materials are multicomponent, multiphasic, and mostly nonequilibrium systems, while living matter is ephemeral. It is thus very difficult to use a thermodynamic approach to predict system behavior. However, at the same time, this makes food systems amenable to kinetic analysis, where researchers try to catch a fleeting glimpse at the dynamics of living processes, structure formation, structure breakdown, state transformations, and the energetics involved in these processes. At the end of the exercise, we talk about free energies of activation, partition functions, orders of processes, and rate constants instead of free energy, enthalpy, and entropy of standard states.

The complexity of a food system might discourage some; however, many of these complex systems are amenable to "coarse-graining," which means that structural or functional averages of ensembles of molecules, higher-order structural units, or entire metabolic pathways can be used to model whole system behavior – a mean field approach. Engineers have known this for a while and carry out phenomenological finite-element and finite-difference analysis; however, there are many approaches available, limited only to the imagination and skill of the researcher. Here is where art meets science, where creativity is key. How do we succeed in describing the behavior of a complex system?

Should we use a deterministic or a stochastic approach? Deterministic processes (cause-effect) follow an exact mathematical rule, while in stochastic processes, the time evolution of a system is represented by a variable whose change is subject to random variation. Deterministic modeling utilizes the tools of differential equations, assuming that the time evolution of a system can be modeled exactly. According to Laplace, "The state of the world at a given instant is defined by an infinite number of parameters, subject to an infinite number of differential equations. If some 'universal' mind could write down all these equations and integrate them, it could then predict with complete exactness the entire evolution of the world in the infinite future." Features common to all deterministic mathematical formulations include that the system is completely defined by mathematical entity ω (can be a set of real numbers, one or more functions, etc.) and that values of ω in the future $t > t_o$ are uniquely determined by $\omega_{o,}$ the initial state of ω, namely, $\omega = F(t_o,\omega_o,t)$. For phenomena described by differential equations, the process of finding their solution involves the integration of these differential equations with the initial conditions $\omega = \omega_o$ and $t = t_o$. Such formulations are exact and direct expressions of the deterministic character of the actual phenomena, of the physical principle of causation. However, this quantitative infinity is coarse compared to the qualitatively inexhaustible character of the real world, and this approach sometimes does not adequately represent the infinite complexity of actual events.

Stochastic modeling, on the other hand, uses probabilities to analyze the time evolution of a system. A Markov chain is probably the most famous example of stochastic modeling, aside from random walks. In a time-discrete Markov process, the probability of each event depends only on the state attained in the previous event in time. Interestingly, most deterministic modeling deals with nonlinear functions, while probabilistic modeling relies on first- and higher-order polynomial functions. This is a consequence of the mathematics used in either approach. Is the world nonlinear and deterministic or linear and probabilistic?

Regardless of the approach, the art of an investigation consists in finding a very simple space Ω (i.e., a set of values of ω or different possible states of the system) such that if we replace the actual process by varying the point ω in a determinate way over this space, we can include all the *essential* aspects of the actual process. In the words of *J.W. Gibbs*, "One of the principal objects of theoretical research in any department of knowledge is to find the point of view from which the subject appears in its greatest simplicity."

At this point, we should discuss other considerations in the modeling process. Are we going to adopt a reductionist view or a wholistic view of the system? A reductionist approach would attempt to describe the behavior of a system from the behavior of the smallest structural unit in that system. In this bottom-up approach, whole system behavior can be predicted from the behavior of ensembles of the smallest structural units. A wholistic approach would seek to describe the behavior of the whole system as a unit, separate from its individual building blocks. In this top-down approach, correlations would be established between macroscopic behavior and particular length scales in a system in an attempt to correlate function to biological or physical structure and their interactions. The approach could also be mechanistic or phenomenological in nature. Mechanistic approaches are based on theory that describes the physical and chemical behavior of a system. A phenomenological approach just utilizes a convenient function which describes system behavior without addressing the reason for that behavior. Finally, I would like to mention the difference between interpolation and extrapolation. A simple phenomenological modeling exercise such as obtaining a least squares estimate of a straight line by linear regression generates a function which can only predict unknown values within the bounds of the data set. It should never be used outside these limits.

This is interpolation. A more sophisticated mechanistic modeling exercise should be able to predict unknown states of the system beyond the data set used to create such model. This is true prediction and constitutes extrapolation from the experimental data set.

In my experience, the modeling process of a complex system involves the following necessary steps for maximum impact:

1. *Write down all the variables that you believe affect the behavior of your system.*
2. *Design the experiment properly, with particular attention to replication. The experiment has to have at least three determinations of three replicates. They are not the same thing!*
3. *Try to control or monitor as many variables as possible during experiments.*
4. *Carry out proper statistical analysis. Any proposed effects or trends need to be statistically significant at least at the 5% level. Determine which variables significantly affect the system and which ones do not;* i.e.*, is the variability attributed to a treatment greater than the one attributed to random experimental error?*
5. *Model the system's behavior using a mechanistic or phenomenological model and test its validity.*
6. *Attribute mechanistic significance to the parameters derived from the model.*

Any discussion about modeling eventually digresses into the realm of the philosophy of carrying out research. Research is the systematic search for new knowledge, while scientific research takes place when research involves a testing step of carefully formulated ideas. Types of research include (a) fact finding, information gathering not intended to derive generalizations or solve problems; (b) critical interpretation, arriving at conclusions through logical reasoning (critical reviews), which is not scientific research; and (c) complete research, solving problems and arriving at generalizations after a thorough search for existent pertinent facts, analysis and logical classification of these facts, and development of a reasonable case using inductive reasoning. The scientific method is a form of complete research plus the development of a hypothesis and controlled experimentation. An essential element of the scientific method includes model building, which has to do with the creation of a hypothesis. Models are central to the scientific method and complete research. Models allow us to explain and create understanding and maybe even predict behavior. These days there is way too much information and not enough knowledge created in research. Moreover, nobody has time to look back and possibly accumulate wisdom (and share it) gathered from new and old knowledge.

This book takes a "stream of consciousness" approach to the subject and is not meant to be exhaustive by any stretch of the imagination. Section 1 is an introduction to the topic of kinetic analysis assuming no prior knowledge of the subject. Section 2 includes nine examples of kinetic modeling that illustrate many techniques and approaches to problems such as color loss, oil migration, biodiesel manufacture, crystallization, nucleation, and protein aggregation. The examples provide a practical handle on how to carry out kinetic analysis, while each chapter is designed to cover a major mathematical technique such as Fourier series analysis and solutions to coupled ordinary differential equations. Examples include both phenomenological and mechanistic modeling. The book is also meant to be a standard reference in the use of the specific models presented for research. The student of modeling should always remember that a model is "entertained" since it is just an attempt to catch a glimpse at the workings of nature. Experimental evidence must always be gathered, which will support or not support the model in the long term. A model is a sophisticated hypothesis.

I finish this brief introduction with two of my favorite quotations which are very relevant to modeling and scientific research, one by a famous scientist and one by a famous wilderness adventurer:

> If nature were not beautiful it would not be worth knowing, and if nature were not worth knowing life would not be worth living. (Henri Poincare)

> Only those who have had the experience can know what a sense of physical and spiritual excitement comes to one who turns his face away from men and towards the unknown. (Bill Mason)

Alejandro G. Marangoni
Guelph, ON, Canada
November 1, 2016

Contents

Part I Kinetic Modelling Basics

Part II Kinetic Modeling of Complex Processes in Food Systems

Part I
Kinetic Modelling Basics

Chapter 1
Fundamentals of Kinetics

1.1 Generalities

The laws of chemical interactions are of great interest to chemists for theoretical and practical reasons. The theories that have been proposed to explain such interactions are empirical, that is, they are based on observable experimental results. Two main approaches have been used to study chemical reactivity, the *thermodynamic* and the *kinetic* approach. In thermodynamics, conclusions are reached on the basis of changes in enthalpy, free energy and entropy that accompany a particular change of state in a system. From the magnitude and sign of the free energy change of a reaction, it is possible to predict the direction in which a chemical change will take place. Thermodynamic quantities do not, however, provide any information on the rate or mechanism of a reaction. These lie in the realm of kinetics. Theoretical analysis of the kinetics, or time course, of processes can also provide valuable information concerning the underlying mechanisms responsible for these processes.

In the study of kinetics, it is necessary to construct a mathematical model that embodies the hypothesized mechanisms. A model is a simplification of the physical world. It is a mathematical representation of observed phenomena. A model is not judged by its scientific accuracy but by its utility. A model may be scientifically untrue yet retained by scientists as long as it adequately describes physical phenomena. A model is “entertained” (as opposed to being “accepted”). A scientist must rigorously test a model and be prepared to discard it and adopt a new one in view of conflicting evidence. The model may be entertained or discarded by examining whether or not it is consistent with the experimental data.

The constructed models must be in accordance with the *law of parsimony* or the *law of succinctness*, also known figuratively as Occam’s Razor. Occam’s Razor is a maxim formulated by the fourteenth -century English logician William of Ockham. The law states that, in the formulation of a theory or a hypothesis to explain natural phenomena, it is necessary to make as few assumptions as possible. Furthermore, assumptions are only to be included if their inclusion results in an observable

A.G. Marangoni, *Kinetic Analysis of Food Systems*,
DOI 10.1007/978-3-319-51292-1_1

difference in the predictions of the theory. While the law of parsimony is one of the fundamental principles of the philosophy of science, it is difficult to apply to food systems. Reactions in foods are rarely so simple that they can be described appropriately with a simple model and a minimum of assumptions.

Modeling consists of two main steps: (1) Formulation of a model and (2) Testing of the model. A *conjecture* about the possible mechanism of the process to be modeled is formulated. An *experiment* is then designed to collect data and test the proposed model. *Analysis* of the data will reveal whether or not the data collected is in accordance with the model. Analysis will provide the grounds to either entertain a model or reject it outright. Model building is a dynamic process. If a model is rejected, a new model is formulated, repeating the entire modeling process.

Kinetic modeling finds many applications in the chemical sciences. In food science, it has been used to study and quantify phenomena such as enzyme catalysis in biotechnological processes, inactivation of microbes and degradative enzymes in food processing, loss of plant pigments, the rate and inhibition of the Maillard reaction, the transesterification of oils, starch degradation and many others. Kinetics is particularly useful in studying the stability of foods over time.

Kinetic modeling is an important tool in food science. Knowing the rate of a reaction given the reactant concentration and its dependence on temperature and other variables can allow for the prediction of the time required for a change in quality to occur. Such predictive power is of great importance to food manufacturers since it allows for the estimation of the shelf-life of a food. As well, studying the dependence of degradative reactions on certain kinetic parameters will provide the framework for the control of these parameters. Such control can enable food processors to decrease the rates of degradative reactions, extending the shelf-life and increasing the quality of foods.

1.2 Basic Definitions

Consider the simple reaction $A \rightarrow B$. The *law of mass action* states that the rate at which the reactant A is converted to product B is proportional to the number of molecules of A available to participate in the chemical reaction. Doubling the concentration of A will double the number of collisions between molecules, leading to an increased rate of product formation.

The *molecularity* of a reaction is the number of reactant molecules participating in a simple reaction consisting of a single elementary step. Reactions can be unimolecular, bimolecular, and termolecular. Unimolecular reactions can include isomerizations ($A \rightarrow B$), and decompositions ($A \rightarrow B + C$). Bimolecular reactions include association ($A + B \rightarrow AB$; $2A \rightarrow A_2$), and exchange reactions ($A + B \rightarrow C + D$ or $2A \rightarrow C + D$). The less common termolecular reactions can also take place ($A + B + C \rightarrow D$). The term molecularity only applies to processes that take place in a single elementary step. Thus, this implies a theoretical understanding of the mechanism of a process. Reactions that take place in a single elementary step are rare. Thus, for more complex reactions it is necessary to specify the molecularity of each elementary step within a multi-step reaction.

1.2.1 The Rate Equation

The *rate equation* or *rate law* is a quantitative expression of the change in concentration of reactant or product molecules in time. It is based on the law of mass action. Writing rate equations that embody the proposed mechanisms is usually the first step in the formulation of a kinetic model. The reaction rate can be obtained from the slope of the *concentration-time* curve for the disappearance of reactant(s) or the appearance of product(s). For example, consider the reaction $\mathrm{A} + 3\mathrm{B} \rightarrow 2\mathrm{C}$. The rate of this reaction can be expressed as either the disappearance of reactant, or the formation of product:

$$rate = -\frac{d[\mathrm{A}]}{dt} = -\frac{1}{3}\frac{d[\mathrm{B}]}{dt} = \frac{1}{2}\frac{d[\mathrm{C}]}{dt} \tag{1.1}$$

The *stoichiometry* of a reaction is the simplest ratio of the number of reactant molecules to the number of product molecules. For example, consider the reaction above. The stoichiometry tells us that one unit of reactant A is needed along with 3 units of reactant B to form 2 units of reactant C. This implies that the rate at which B is consumed is three times the rate at which A is consumed. Likewise, this tells us that the rate at which C is formed is twice the rate at which A is consumed.

Empirically, one also finds that the rate of a reaction is proportional to the amount of reactant present, raised to an exponent n:

$$rate \propto [\mathrm{A}]^n \tag{1.2}$$

where n is the *order* of the reaction. Order is an *empirical* quantity, that is, it is mathematically consistent with observed phenomena in the physical world.

If the rate of a reaction is *independent* of a particular reactant concentration, the reaction is considered to be *zero order* with respect to the concentration of that reactant ($n = 0$). If the rate of a reaction is *directly proportional* to a particular reactant concentration, the reaction is considered to be *first order* with respect to the concentration of that reactant ($n = 1$). If the rate of a reaction is proportional to the square of a particular reactant concentration, the reaction is considered to be *second order* with respect to the concentration of that reactant ($n = 2$). In general, for any reaction $A + B + C + \ldots \rightarrow P$, the rate equation can be generalized as:

$$Rate \propto [\mathrm{A}]^a [\mathrm{B}]^b [\mathrm{C}]^c \ldots \tag{1.3}$$

where the exponents a, b, c correspond, respectively, to the order of the reaction with respect to reactants A, B, and C.

The rate equation for the proportionality given in Eq. 1.2 can be re-written as:

$$rate = -\frac{d[\mathrm{A}]}{dt} = k_r [\mathrm{A}]^n \tag{1.4}$$

where k_r is the rate constant of the reaction.

The *rate constant* (k_r) of a reaction is a concentration-independent measure of the velocity of a reaction. For a first order reaction, k_r has units of $(\text{time})^{-1}$, while for a second order reaction, k_r has units of $(\text{concentration})^{-1}(\text{time})^{-1}$. In general, the rate constant of an nth order reaction has units of $(\text{concentration})^{-(n-1)}(\text{time})^{-1}$.

As implicitly stated above, the rate of a reaction can be obtained from the slope of the concentration-time curve for the disappearance of reactant(s) or the appearance of product(s). Typical reactant concentration-time curves for zero, first, second and third order reactions are shown in Fig. 1.1. The dependence of the different reaction rates or velocities on reactant concentration are shown in Fig. 1.2.

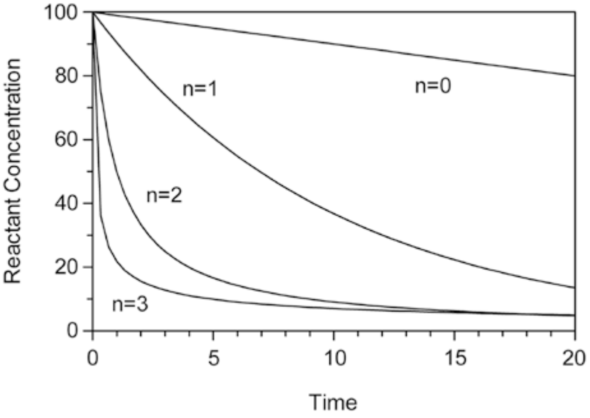


Fig. 1.1 Changes in reactant concentration as a function of time for zero-, first-, second- and third-order reactions

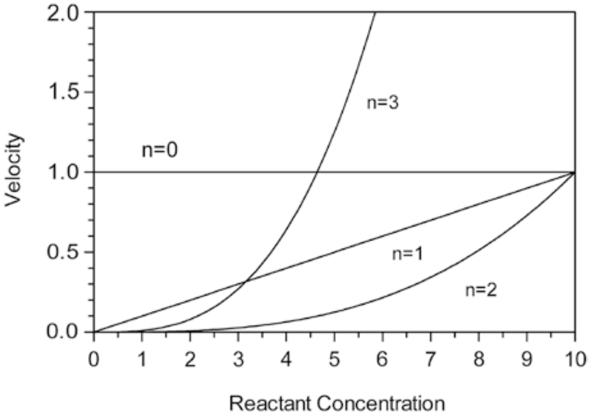


Fig. 1.2 Changes in reaction velocity as a function of reactant concentration for zero-, first-, second-, and third-order reactions

1.2.2 Integrated Rate Equations

By integrating the rate equations, it is possible to obtain analytical expressions that describe changes in the concentration of reactants or products as a function of time. As will be described in the next section, these integrated rate equations are extremely useful in the experimental determination of rate constants and reaction order.

1.2.2.1 Zero Order Integrated Rate Equation

The rate equation for a zero order reaction can be expressed as:

$$rate = \frac{d[\mathrm{A}]}{dt} = -k_r [\mathrm{A}]^0 \tag{1.5}$$

Since $[\mathrm{A}]^0 = 1$, integration of the above equation for the boundary conditions $A = A_o$ at $t = 0$ and $A = A_t$ at time t,

$$\int_{A_o}^{A_t} d[\mathrm{A}] = -k_r \int_0^t dt \tag{1.6}$$

yields the integrated rate equation for a zero order reaction:

$$[A_t] = [A_o] - k_r t \tag{1.7}$$

where $[A_t]$ is the concentration of reactant A at time t, and $[A_o]$ is the initial concentration of reactant A at $t = 0$. For a zero order reaction, a plot of $[A_t]$ versus time yields a straight line with slope $-k_r$ (Fig. 1.3).

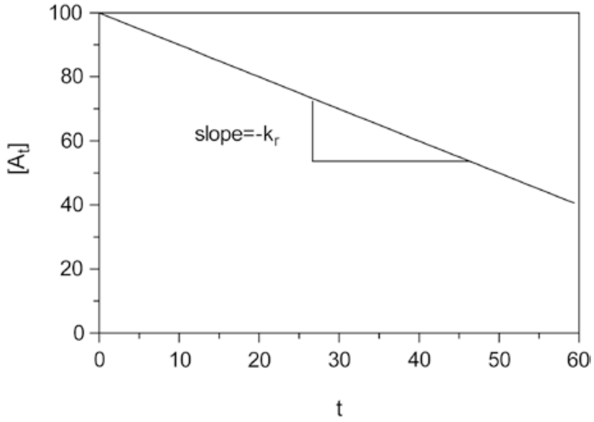


Fig. 1.3 Determination of the reaction rate constant (k_r) for a zero-order reaction from a graph showing the change in reactant concentration as function of time

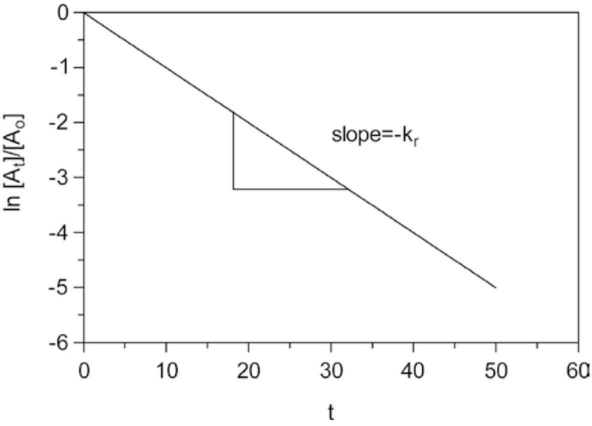


Fig. 1.4 Determination of the reaction rate constant (k_r) for a first-order reaction from a semi-logarithmic graph showing the change in reactant concentration as a function of time

1.2.2.2 First Order Integrated Rate Equation

The rate equation for a first order reaction can be expressed as:

$$rate = \frac{d[A]}{dt} = -k_r [A] \tag{1.8}$$

Integration of the above equation for the boundary conditions $A = A_o$ at $t = 0$ and $A = A_t$ at time t,

$$\int_{A_o}^{A_t} \frac{d[A]}{[A]} = -k_r \int_0^t dt \tag{1.9}$$

yields the integrated rate equation for a first order reaction:

$$\ln \frac{A_t}{A_o} = -k_r t \tag{1.10}$$

This can be re-arranged to give:

$$[A_t] = [A_o] e^{-k_r t} \tag{1.11}$$

For a first order reaction, a plot of $\ln \frac{[A_t]}{[A_o]}$ versus time yields a straight line with slope $-k_r$ (Fig. 1.4).

1.2.2.3 Second Order Integrated Rate Equation

The rate equation for a second order reaction can be expressed as:

$$rate = \frac{d[A]}{dt} = -k_r [A]^2 \tag{1.12}$$

Integration of the above equation for the boundary conditions $A = A_o$ at $t = 0$ and $A = A_t$ at time t,

$$\int_{A_o}^{A_t} \frac{d[A]}{[A]^2} = -k_r \int_0^t dt \tag{1.13}$$

yields the integrated rate equation for a second order reaction:

$$\frac{1}{[A_t]} = \frac{1}{[A_o]} + k_r t \tag{1.14}$$

The following equation can be re-arranged to give:

$$[A_t] = \frac{[A_o]}{1 + [A_o] k_r t} \tag{1.15}$$

For a second order reaction, a plot of $\frac{1}{[A_t]}$ against t yields a straight line with positive slope k_r (Fig. 1.5).

For the second-order bimolecular reaction $A + B \rightarrow P$, it is possible to express the rate of the reaction in terms of the amount of reactant that is converted to product (P) in time:

$$\frac{d[P]}{dt} = k_r [A_o - P][B_o - P] \tag{1.16}$$

Integration of the above equation using the method of partial fractions for the boundary conditions $A = A_o$ and $B = B_o$ at $t = 0$ and $A = A_t$ and $B = B_t$ at time t,

$$\frac{1}{([A_o] - [B_o])} \int_0^{P_t} \left(\frac{dP}{[B_o - P]} - \frac{dP}{[A_o - P]} \right) = -k_r \int_0^t dt \tag{1.17}$$

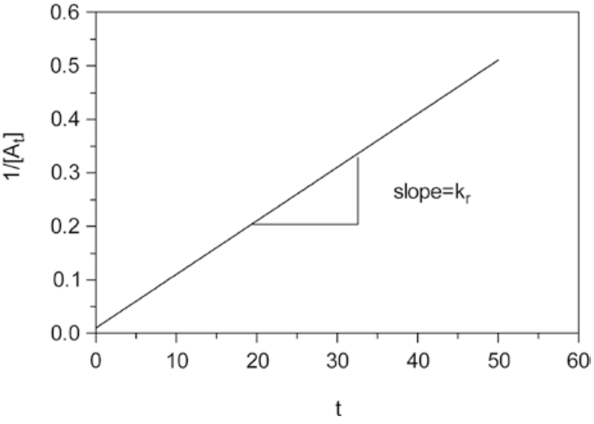


Fig. 1.5 Determination of the reaction rate constant (k_r) for a second-order reaction from a linear plot of the reactant concentration as a function of time

yields the integrated rate equation for a second order reaction in which two different reactants participate:

$$\frac{1}{\left[A_o - B_o\right]} \ln \frac{\left[B_o\right]\left[A_t\right]}{\left[A_o\right]\left[B_t\right]} = k_r t \tag{1.18}$$

where $[A_t] = [A_o - P_t]$ and $[B_t] = [B_o - P_t]$. For this type of second order reaction, a plot of $\frac{1}{\left[A_o - B_o\right]} \ln \frac{B_o A_t}{A_o B_t}$ versus time yields a straight line with positive slope k_r.

1.2.2.4 Higher Order Reactions

For any reaction where n>1, the integrated rate equation has the general form:

$$\frac{1}{(n-1)\left[A_t\right]^{(n-1)}} = \frac{1}{(n-1)\left[A_o\right]^{(n-1)}} + k_r t \tag{1.19}$$

Or

$$\left[A_t\right] = \frac{\left[A_o\right]}{\sqrt[(n-1)]{1 + (n-1)\left[A_o\right]^{(n-1)} k_r t}} \tag{1.20}$$

For an nth order reaction, a plot of $1/[(n-1)\,[A_t]^{(n-1)}]$ versus time yields a straight line with positive slope k_r.

1.2.2.5 Opposing Reactions

For the simplest case of an opposing reaction, $A \rightleftharpoons B$,

$$\frac{d[A]}{dt} = -k_1[A] + k_{-1}[B] \tag{1.21}$$

where k_1 and k_{-1} represent, respectively, the rate constant for the forward $A \rightarrow B$ and reverse $B \rightarrow A$ reactions. It is possible to express the rate of the reaction in terms of the amount of reactant that is converted to the product B in time (Fig. 1.6):

$$\frac{d[B]}{dt} = k_1\left[A_o - B\right] - k_{-1}[B] \tag{1.22}$$

At equilibrium, $\frac{d[B]}{dt} = 0$, and $[B] = [B_e]$, and it is therefore possible to obtain expression for k_{-1} and $k_1[A_o]$:

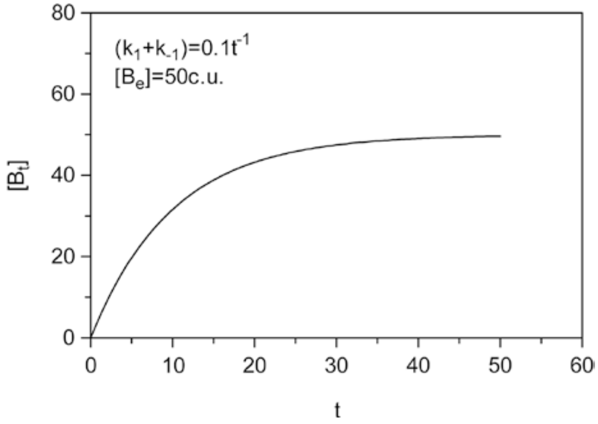


Fig. 1.6 Changes in production concentration as a function of time for a reversible reaction of the form A $\rightleftharpoons$B

$$k_{-1} = \frac{k_1[A_o - B_e]}{[B_e]} \quad and \quad k_1[A_o] = (k_{-1} + k_1)[B_e] \tag{1.23}$$

Substituting $\frac{k_1[A_o - B_e]}{[B_e]}$ for k_1 into the rate equation, we obtain:

$$\frac{d[B]}{dt} = k_1[A_o - B] - \frac{k_1[A_o - B_e][B]}{B_e} \tag{1.24}$$

Summing together the terms on the right side of the equation, substituting $(k_{-1}+k_1)[B_e]$ for $k_1[A_o]$, and integrating for the boundary conditions $B=0$ at $t=0$ and $B=B_t$ at time t,

$$\int_0^{B_t} \frac{dB}{\left(\frac{[B_e - B]}{[B_e]}\right)} = (k_1 + k_{-1})\int_0^t dt \tag{1.25}$$

yields the integrated rate equation for the opposing reaction $A \rightleftharpoons B$:

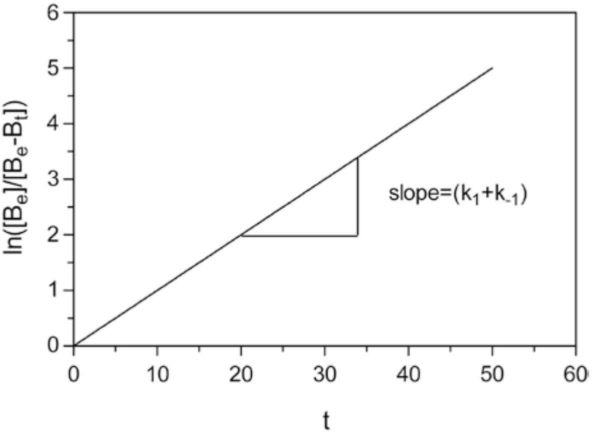
$$\ln\frac{[B_e]}{[B_e - B_t]} = (k_1 + k_{-1})t \tag{1.26}$$

The following equation can be re-arranged to yield:

$$[B_t] = [B_e] - [B_e]e^{-(k_1+k_{-1})t} \tag{1.27}$$

A plot of $\ln\left(\frac{[B_e]}{[B_e - B]}\right)$ versus time results in a straight line with positive slope (k_1+k_{-1}) (Fig. 1.7).

Fig. 1.7 Linear plot of changes in product concentration as a function of time used in the determination of forward (k_1) and reverse (k_{-1}) reaction rate constants



1.2.2.6 Reaction Half-life

The half-life is another useful measure of the rate of a reaction. A reaction half-life is the time required for the concentration of the initial reactant(s) to decrease by one-half. Useful relationships between the rate constant and the half-life can be derived using the integrated rate equations by substituting $\frac{1}{2}A_0$ for A_t.

The resulting expressions for the half-life of reactions of different orders (n) are as follows:

$$n = 0 \ldots t_{\frac{1}{2}} = \frac{0.5[A_0]}{k_r} \tag{1.28}$$

$$n = 1 \ldots t_{\frac{1}{2}} = \frac{\ln(2)}{k_r} \tag{1.29}$$

$$n = 2 \ldots t_{\frac{1}{2}} = \frac{1}{k_r[A_0]} \tag{1.30}$$

$$n = 3 \ldots t_{\frac{1}{2}} = \frac{3}{2k_r[A_0]^2} \tag{1.31}$$

The half-life of an nth order reaction where $n > 1$ can be calculated from the following expression:

$$n = 2 \ldots t_{\frac{1}{2}} = \frac{1-(0.5)^{(n-1)}}{(n-1)k_r[A_0]^{(n-1)}} \tag{1.32}$$

1.3 Dependence of Reaction Rates on Temperature

1.3.1 Theoretical Considerations

The rates of chemical reactions are highly dependent on temperature. Temperature affects the rate constant of a reaction, but not the order of the reaction. Classic thermodynamic arguments will be used to derive an expression for the relationship between the reaction rate and temperature.

This treatment involves knowledge of the concept of molar standard state, enthalpy ΔH^0, entropy ΔS^0 and free energy ΔG^0 changes of the system. These thermodynamic functions have the form:

$$\Delta H^0 = \int_{298K}^{T} C_p dT \tag{1.33}$$

$$\Delta S^0 = \int_{298K}^{T} \frac{C_p}{T} dT \tag{1.34}$$

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 \tag{1.35}$$

where the standard state enthalpy, entropy and free energy changes in the system are always between the reference state temperature of 298 K and a given temperature T. The molar heat capacity of the system (C_p) can be constant, or variable, in the temperature range (298 K, T). These thermodynamic functions are only dependent on temperature and correspond to a standard state where the pressure is 1 atm.

The molar standard state free energy change of a reaction at *equilibrium* is a function of the equilibrium constant (K), and is related to changes in the molar standard state enthalpy and entropy as described by the Gibbs-Helmholtz equation:

$$\Delta G^0 = -RT \ln K \tag{1.36}$$

Re-arrangement of the above equation and substitution of ΔG^0 with $\Delta H^0 - T\Delta S^0$ yields the well-known van't Hoff equation:

$$\ln K = -\frac{\Delta H^o}{RT} + \frac{\Delta S^o}{R} \tag{1.37}$$

The change in ΔS^0 due to a temperature change from T_1 to T_2 is given by:

$$\Delta S^0_{T_2} = \Delta S^0_{T_1} + \Delta C_p \ln \frac{T_2}{T_1} \tag{1.38}$$

While the change in ΔH^0 due to a temperature change from T_1 to T_2 is given by:

$$\Delta H^0_{T_2} = \Delta H^0_{T_1} + \Delta C_p \left(T_2 - T_1\right) \tag{1.39}$$

If the standard heat capacities of reactants and products are the same i.e., $\Delta C_p = 0$, then ΔS^0 and ΔH^0 are independent of temperature. Subject to the condition that the difference in the heat capacities between reactants and products is zero, differentiation of Eq. 1.37 with respect to temperature yields a more familiar form of the van't Hoff equation:

$$\frac{d\ln K}{dT} = \frac{\Delta H^0}{RT^2} \tag{1.40}$$

For an endothermic reaction, ΔH^0 is positive, while for an exothermic reaction ΔH^0 is negative. The van't Hoff equation predicts that ΔH^0 of a reaction defines the effect of temperature on the equilibrium constant. For an endothermic reaction, K increases as T increases, while for an exothermic reaction, K decreases as T increases. These predictions are in agreement with le Chatelier's principle, which states that increasing the temperature of an equilibrium reaction mixture causes the reaction to proceed in the direction that absorbs heat. The van't Hoff equation is used for the determination of the ΔH^0 of a reaction by plotting ln K against $\frac{1}{T}$. The slope of resulting line corresponds to $\frac{-\Delta H^0}{R}$ (Fig. 1.8). It is also possible to determine the ΔS^0 of the reaction from the y-intercept, which corresponds to $\frac{\Delta S^0}{R}$. It is important to reiterate that this treatment applies only for cases where the heat capacities of the reactants and products are equal and temperature-independent.

Enthalpy changes are related to changes in internal energy:

$$\Delta H^o = \Delta E^o + \Delta\left(PV\right) = \Delta E^o + P_1V_1 - P_2V_2 \tag{1.41}$$

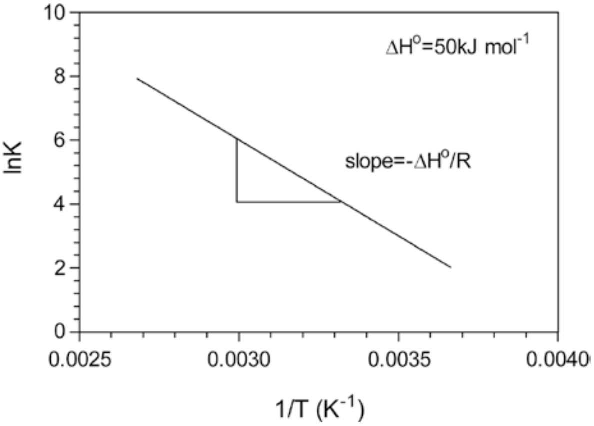


Fig. 1.8 A van't Hoff plot used in the determination of the standard-state enthalpy ΔH^0 of a reaction

Hence, ΔH^0 and ΔE^0 differ only by the difference in the PV product of the final and initial states. For a chemical reaction at constant pressure in which only solids and liquids are involved, $\Delta PV \approx 0$ and therefore ΔH^0 and ΔE^0 are nearly equal. For gas phase reactions, $\Delta PV \neq 0$, unless the number of moles of reactants and products remains the same. For ideal gases, it can easily be shown that $\Delta PV = (\Delta n)RT$. Thus, for gas-phase reactions, if $\Delta n = 0$, then $\Delta H^0 = \Delta E^0$.

At equilibrium, the rate of the forward reaction (v_1) is equal to the rate of the reverse reaction (v_{-1}), $v_1 = v_{-1}$. Therefore, for the reaction A↔B at equilibrium, $k_1[A_e] = k_{-1}[B_e]$

and therefore:

$$K = \frac{[\text{Products}]}{[\text{Reactants}]} = \frac{[B_e]}{[A_e]} = \frac{k_1}{k_{-1}} \tag{1.42}$$

Considering the above, the van’t Hoff equation can therefore be rewritten as:

$$\frac{d\ln k_1}{dT} - \frac{d\ln k_{-1}}{dT} = \frac{\Delta E^o}{RT^2} \tag{1.43}$$

The change in the standard state internal energy of a system undergoing a chemical reaction from reactants to products ΔE^0 is equal to the energy required for reactants to be converted to products minus the energy required for products to be converted to reactants (Fig. 1.9). Moreover, the energy required for reactants to be converted to products is equal to the difference in energy between the ground state and transition state of the reactants ($\Delta E_1^{\#}$), while the energy required for products to be converted to reactants is equal to the difference in energy between the ground state and transition state of the products ($\Delta E_1^{\#}$). Therefore, the change in the internal energy of a system undergoing a chemical reaction from reactants to products can be expressed as:

$$\Delta E^o = E_{products} - E_{reactants} = \Delta E_1^{\#} - \Delta E_{-1}^{\#} \tag{1.44}$$

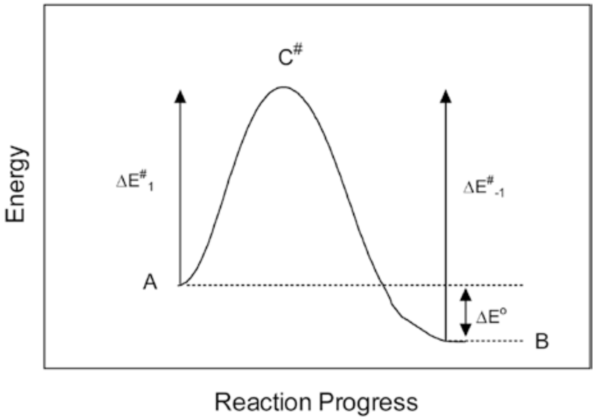


Fig. 1.9 Changes in the free energy of a system undergoing a chemical reaction from substrate A to product B. $\Delta E^{\ddagger}$ corresponds to the energy barrier/energy of activation for the forward (1) and reverse (−1) reactions, $C^{\ddagger}$corresponds to the transition state structure and $\Delta E^{\ddagger}$ corresponds to the standard-state difference in the internal energy between products and reactants

Equation 1.43 can therefore be expressed as two separate differential equations corresponding to the forward and reverse reactions namely:

$$\frac{dlnk_1}{dT} = \frac{\Delta E_1^{\#}}{RT^2} + C \tag{1.45a}$$

and

$$\frac{dlnk_{-1}}{dT} = \frac{\Delta E_{-1}^{\#}}{RT^2} + C \tag{1.45b}$$

Arrhenius determined that for many reactions, $C = 0$. Therefore, indefinite integration of the above equation for either the forward or reverse reactions,

$$\int d \ln k_r = \frac{\Delta E^{\#}}{R} \int \frac{dT}{T^2} \tag{1.46}$$

yields an expression for the relationship between the rate constant of a reaction and temperature. Upon re-arrangement, the familiar Arrhenius equation is obtained

$$\ln k_r = \ln \Delta A - \frac{\Delta E^{\#}}{RT} \text{ or } k_r = Ae^{-\frac{\Delta E^{\#}}{RT}} \tag{1.47}$$

$\Delta E^{\#}$, or E_a as Arrhenius defined this term, is the energy of activation for a chemical reaction, while A is called the *Arrhenius pre-exponential factor*. Since, in first-order kinetics, A has units of (*time*)$^{-1}$, it is also sometimes called the *frequency factor*. The Arrhenius pre-exponential factor has the same dimensions as the rate constant and is related to the frequency of collisions between reactant molecules.

1.3.2 Energy of Activation

Figure 1.9 depicts a potential energy reaction coordinate for a hypothetical reaction A↔B. In order for A molecules to be converted to B (forward reaction), or for B molecules to be converted to A (reverse reaction), they must acquire energy to form an activated complex $C^{\#}$. This potential energy barrier is therefore called the *energy of activation* of the reaction. This energy of activation is the minimum energy that must be acquired by the system's molecules in order for the reaction to take place. Only a small fraction of the molecules may possess sufficient energy to react. The rate of the forward reaction depends on $\Delta E_1^{\#}$, while the rate of the reverse reaction depends on $\Delta E_{-1}^{\#}$ (Fig. 1.9). As will be shown later, the rate constant is inversely proportional to the energy of activation.

In order to determine the energy of activation of a reaction, it is necessary to measure the rate constant of a particular reaction at different temperatures. A plot of $\ln k_r$ versus $\frac{1}{T}$ yields a straight line with slope $\frac{-\Delta E^{\#}}{R}$ (Fig. 1.10).

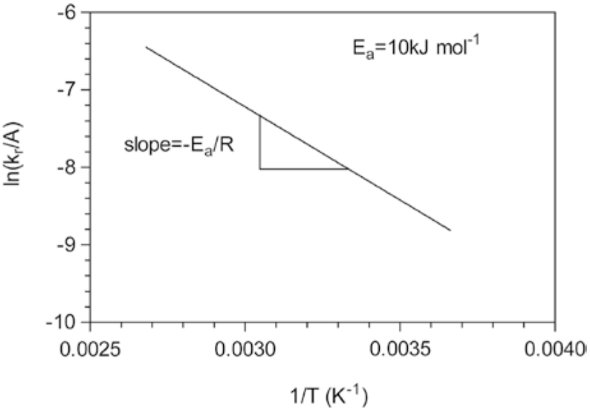


Fig. 1.10 An Arrhenius plot used in the determination of the energy of activation E_a of a reaction

Alternatively, but not recommended, is the integration of Eq. 1.46 as a definite integral with appropriate boundary conditions,

$$\int_{k_1}^{k_2} d\,ln\,k_r = \frac{\Delta E^{\#}}{R}\int_{T_1}^{T_2}\frac{dT}{T^2} \tag{1.48}$$

which yields the following expression:

$$\ln\frac{k_2}{k_1} = \frac{\Delta E^{\#}}{R}\left(\frac{T_2 - T_1}{T_2 T_1}\right) \tag{1.49}$$

This equation can be used to obtain the energy of activation, or predict the value of the rate constant at T_2 from knowledge of the value of the rate constant at T_1, and of $\Delta E^{\#}$.

1.4 Theory of Reaction Rates

Absolute reaction rate theory will be discussed briefly in this section. Collision theory will not be explicitly developed since it is less applicable to the complex systems studied. Absolute reaction rate theory is a collision theory that assumes that chemical activation occurs through collisions between molecules. The central postulate of this theory is that the rate of a chemical reaction is given by the rate of passage of the activated complex through the transition state.

This theory is based on two assumptions, a dynamical bottleneck assumption and an equilibrium assumption. The first asserts that the rate of a reaction is controlled by the decomposition of an activated transition-state complex, and the second one asserts that an equilibrium exists between reactants (A and B) and the transition-state complex, $C^{\#}$, namely $A+B \rightleftarrows C^{\#} \rightarrow C+D$.

It is therefore possible to define an equilibrium constant for the conversion of reactants in the ground state into an activated complex in the transition state. For the above reaction:

$$K^{\#} = \frac{\left[C^{\#}\right]}{[A][B]} \tag{1.50}$$

As discussed previously, $\Delta G^o = -RT\ ln\ K$, and $\ln \mathrm{K} = \ln k_1 - \ln k_{-1}$. Thus, in an analogous treatment to the derivation of the Arrhenius equation (see above), it would be straightforward to show that:

$$k_r = ce^{-\frac{\Delta G^{\#}}{RT}} = cK^{\#} \tag{1.51}$$

Where $\Delta G^{\#}$ is the free energy of activation for the conversion of reactants into the activated complex $C^{\#}$. By using statistical thermodynamic arguments, it is possible to show that the constant c equals:

$$c = \kappa \nu \tag{1.52}$$

where κ is the transmission coefficient and ν is the frequency of the normal mode oscillation of the transition state complex along the reaction coordinate; more rigorously, the average frequency of barrier crossing. The transmission coefficient, which can differ dramatically from unity, includes many correction factors, including tunneling, barrier re-crossing correction and solvent frictional effects. The rate of a chemical reaction depends on the equilibrium constant for the conversion of reactants into activated complex.

Since $\Delta G = \Delta H - T\Delta S$, it is possible to rewrite Eq. 1.51 as:

$$k_r = \kappa \nu e^{\frac{\Delta S^{\#}}{R}} e^{-\frac{\Delta H^{\#}}{RT}} \tag{1.53}$$

Consider $\Delta H = \Delta E + (\Delta n)RT$, where Δn equals the difference between the number of moles of activated complex n_{ac} and the moles of reactants (n_r). The term n_r also corresponds to the molecularity of the reaction, e.g., unimolecular, bimolecular. At any particular time, $n_r < < n_{ac}$ and therefore $\Delta H \approx \Delta E - n_r RT$. Substituting this expression for the enthalpy change into Eq. 1.53 and rearranging, we obtain:

$$k_r = \kappa \nu e^{\left(n_r + \frac{\Delta S^{\#}}{R}\right)} e^{-\frac{\Delta E^{\#}}{RT}} \tag{1.54}$$

Comparison of this equation with the Arrhenius equation sheds light on the nature of the frequency factor:

$$A = \kappa \nu e^{\left(n_r + \frac{\Delta S^{\#}}{R}\right)} \tag{1.55}$$

The concept of entropy of activation ($\Delta S^{\#}$) is of utmost importance for the understanding of reactivity. Two reactions with similar ($\Delta E^{\#}$) at the same temperature can proceed at appreciably different rates. This effect is due to differences in their entropies of activation. The entropy of activation corresponds to difference in entropy between the ground state and transition state of the reactants. Recalling that entropy is a measure of the randomness of a system, a positive $\Delta S^{\#}$ suggests that the transition state is more disordered (more degrees of freedom) than the ground state. Alternatively, a negative $\Delta S^{\#}$ suggests that the transition state is more ordered (less degrees of freedom) than the ground state. Freely diffusing, non-interacting molecules have many translational, vibrational and rotational degrees of freedom. When two molecules interact at the onset of a chemical reaction and pass into a more structured transition state, some of these degrees of freedom will be lost. For this reason, most entropies of activation for chemical reactions are negative. When the change in entropy for the formation of the activated complex is small ($\Delta S^{\#} \approx 0$), then the rate of the reaction is solely controlled by the energy of activation ($\Delta E^{\#}$).

It is interesting to use the concept of entropy of activation in order to explain the failure of collision theory to explain reactivity. Consider that for a bimolecular reaction $A + B \rightarrow P$, the frequency factor (A) equals the number of collisions per unit volume between reactant molecules (Z) times a steric, or probability factor (P), namely:

$$A = PZ = \kappa \nu e^{\left(n_r + \frac{\Delta S^{\#}}{R}\right)} \tag{1.56}$$

If only a fraction of the collisions result in conversion of reactants into products, then $P < 1$, implying a negative $\Delta S^{\#}$. For this case, the rate of the reaction will be slower than predicted by collision theory. If a greater number of reactant molecules than predicted from the number of collisions are converted into products, then $P > 1$, implying a positive $\Delta S^{\#}$. For this case, the rate of the reaction will be faster than predicted by collision theory. On the other hand, when $P = 1$ and $\Delta S^{\#}$, predictions from collision theory and absolute rate theory agree.

1.5 Reaction in Solution

The major difference between reactions in the gas phase and in solution is *solvent effects*. In a liquid, molecules are much closer together and the energy difference between reactants and the activated complex $E^{\#}$ should be smaller than in the gas phase. In the gas phase, two molecules that collide with each other do so only once and generally fly away. In solution, however, molecules are trapped in a solvent "cage". This cage effect causes interacting molecules to remain together for longer periods of time, and molecules tend to collide with each other hundreds of times before drifting apart. Thus, for a case of a reaction with a low energy of activation, this cage effect virtually ensures reaction during each encounter. The steric factor is less important here since there is a high probability that the reacting molecules will become properly oriented for reaction during the collisions in the “cage”. Under these conditions, reactions are limited mainly by the diffusion coefficient of the molecules in the medium.

The situation is quite different for charged species in solution. The solvation around charged ions can greatly influence the rate of chemical reactions via an effect on the entropy of activation ($\Delta S^{\#}$). Moreover, steric effects will also greatly influence the sign and magnitude of $\Delta S^{\#}$. The reaction rate is strongly dependent on the ionic strength of the solution. This kinetic salt effect can be modeled using:

$$\log_{10}\frac{k}{k_o} = z_a z_b A\sqrt{I} \tag{1.57}$$

where k and k_0 correspond to the rate constants of a reaction at a particular ionic strength (I) and the case where $I \cong 0$. Recall that ionic strength is defined as $I = \frac{1}{2}\sum_{i=1}^{n} c_i z_i^2$, where $[c]_i$ corresponds to the molar concentration of the ith ion in solution and z_i corresponds to the net charge of that ion. The constant A depends on the nature of the solvent and temperature, while z_a and z_b correspond to the charge of reactants A and B. An increase or decrease in $\frac{k}{k_o}$ will depend on the relative charge on the reactants, and the ionic strength of the medium.

1.6 Diffusion Controlled Reactions

A recurring question in the area of kinetics is how fast reactions can take place in solution. The ultimate limit to the speed at which a reaction can take place is set by the rate of approach of the reacting species, which in turn is set by their rate of diffusion. Smoluchowski showed that the rate of a diffusion controlled reaction (v), where all encounters lead to a productive reaction outcome, is given by:

$$v = k_D N_A N_B \tag{1.58}$$

where N_a and N_B correspond to the number of reacting molecules A and B, and k_D, the diffusion controlled rate constant of the reaction $[cm^3 \cdot molecule^{-1} \cdot s^{-1}]$ is given by:

$$k_D = 4\pi\left(D_A + D_B\right)\left(r_A + r_B\right) \tag{1.59}$$

where D_A and D_B are the diffusion coefficients for molecules A and B $[cm^2 \cdot s^{-1}]$, and r_A and r_B correspond to the radii of the reacting molecules $[cm]$. If we simplify this model for the case where $D_A = D_B$ and $r_A = r_B$, we then obtain:

$$k_D = 16\pi r D \tag{1.60}$$

The Stokes-Einstein equation,

$$D = \frac{k_b T}{6\pi\eta r} \tag{1.61}$$

states that the diffusion coefficient of a spherical molecule corresponds to the ratio of the thermal (or kinetic) energy of the molecule ($k_B T$), where k_B is Boltzman's constant and T is the absolute temperature, to the viscous energy dissipation, or viscous resistance to diffusion ($6\pi\eta r$), where η is the solvent viscosity and r is the radius of the particle. This energy dissipation arises from the frictional forces between the solute molecules and the solvent molecules. The ratio of these opposing quantities defines how easily a molecule diffuses in solution.

Introducing Eq. 1.61 into Eq. 1.60, we obtain and expression of the rate constant of a diffusion controlled reaction as a function of temperature and solvent viscosity, namely:

$$k_D = \frac{8}{3}\frac{k_b T}{\eta} \tag{1.62}$$

Multiplying this expression by Avogadro's number (N_{av}) gives us the rate constant per mole [$\mathrm{cm}^3 \cdot mol^{-1} \cdot s^{-1}$],

$$k_D = \frac{8}{3}\frac{RT}{\eta} \tag{1.63}$$

where R is the universal gas constant ($R = N_{av} \cdot k_B$). Since 1 cm^3 equals 0.001 L, and $1\ \frac{mol}{L}$ is equal to $1M$, Eq. 1.63 can be expressed as:

$$k_D = \frac{8}{3000}\frac{RT}{\eta} \tag{1.64}$$

in units of $M^{-1}\mathrm{s}^{-1}$.

Two characteristics of diffusion controlled reactions are that they have zero activation energy, and the rate is inversely proportional to the viscosity of the medium, which itself has an Arrhenius-like temperature dependence.

1.7 Experimental Determination of Reaction Order and Rate Constants

1.7.1 Differential Method (Initial Rate Method)

Knowledge of the value of the rate of the reaction at different reactant concentrations would allow for the determination of the rate and order of a chemical reaction. For the reaction $A \rightarrow B$, for example, reactant or product concentration-time curves are determined at different initial reactant concentrations. The absolute value of slope of the curve at $t = 0$, $\left|\frac{dA}{dt}\right|_o$ or $\left|\frac{dB}{dt}\right|_o$, corresponds to the initial rate or initial velocity of the reaction (Fig. 1.11).

As shown before, the reaction velocity (v_A) is related to reactant concentration,

$$v_A = \left|\frac{dA}{dt}\right| = k_r\left[A\right]^n \tag{1.65}$$

Taking logarithms on both sides of the above equation results in the expression:

$$\log v_A = \log k_r + n\log\left[A\right] \tag{1.66}$$

A plot of the logarithm of the initial rate against the logarithm of the initial reactant concentration yields a straight line with a y-intercept corresponding to $\log k_r$, and a slope corresponding to n (Fig. 1.12).

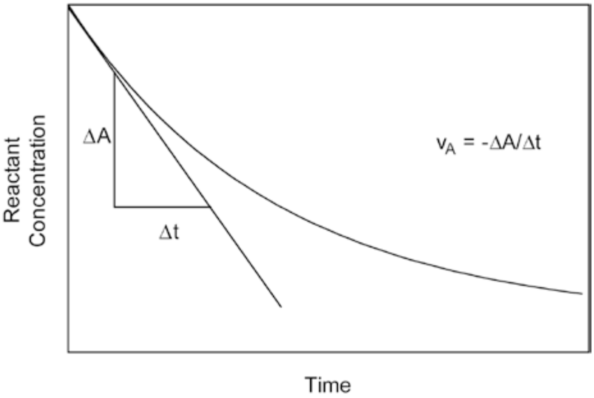


Fig. 1.11 Determination of the initial velocity of a reaction as the instantaneous slope of the substrate depletion curve in the vicinity of t = 0

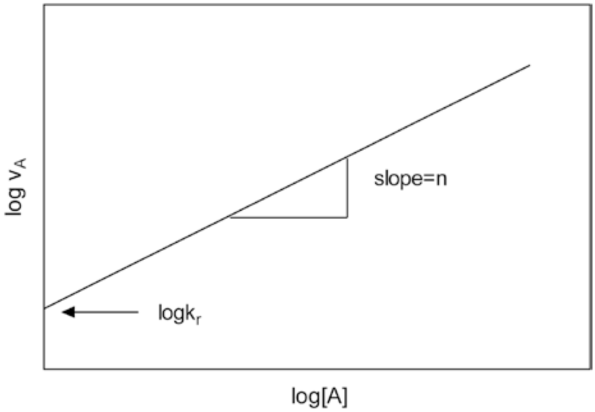


Fig. 1.12 Logarithmic plot of initial velocity versus initial substrate concentration used in the determination of the reaction rate constant (k_r) and the order of the reaction

For more accurate determinations of the initial rate, changes in reactant concentration are measured over a small time period, where less than 1% conversion of reactant to product has taken place.

1.7.2 Integral Method

In this method, the rate constant and order of a reaction are determined from least-squares fits of the integrated rate equations to reactant depletion or product accumulation concentration-time data. At this point, knowledge of the reaction order is required. If the order of the reaction is not known, one is assumed or guessed at, e.g., $n = 1$. If necessary, data is transformed accordingly, e.g., $\ln\left(\frac{A_t}{A_o}\right)$, if a linear first-order model is to be used. The model is then fitted to the data using standard least-squares minimization protocols, i.e., linear or non-linear regression. From this exercise, a "best-fit" slope, y-intercept, their corresponding standard errors, as well as the pearson correlation coefficient, r for the fit, is determined. The r-squared statistic is used as a measure of the fraction of the total variance accounted for by the model. The closer the values of $|r^2|$ to one, the better the fit of the model to the data. This procedure is repeated assuming a different reaction order, e.g., $n=2$. The order of the reaction would thus be determined by comparing the correlation coefficients for the different fits of the kinetic models to the transformed data. The model that fits the data best defines the order of that reaction. The rate constant for the reaction, and its corresponding standard error, is then determined using the appropriate model. If the coefficients are similar, further experimentation may be required to determine the order of the reaction.

The advantage of the differential method over the integral method is that no reaction order needs to be assumed. The reaction order is determined directly from the data analysis. On the other hand, determination of initial rates can be rather inaccurate.

In order to use integrated rate equations, knowledge of reactant or product concentrations is not an absolute requirement. Any parameter *proportional* to reactant or product concentration can be used in the integrated rate equations, e.g., absorbance or transmittance, turbidity, conductivity, pressure, volume, among many others. However, certain modifications may have to be introduced into the rate equations, since reactant concentration, or related parameters, may not decrease to zero --- a minimum, non-zero value (A_{min}) might be reached. For product concentration, and related parameters, a maximum value (P_{max}) may be reached, which does not correspond to 100% conversion of reactant to product. A certain amount of product may even be present at $t=0$ (P_o). The modifications introduced into the rate equations are straightforward. For reactant (A) concentration:

$$[A_t] \rightarrow [A_t - A_{\min}] \; and \; [A_o] \rightarrow [A_o - A_{\min}]$$

For product (P) concentration:

$$[P_t] \to [P_t - P_o]\, and\, [P_o] \to [P_{\max} - P_o]$$

1.8 Modeling Complex Reaction Pathways

In this section, we will briefly discuss strategies for tackling more complex reaction mechanisms. The first step in any kinetic modeling exercise is to write down the differential equations and mass balance that describe the process.

Consider the reaction $A \xrightarrow{k_1} B \xrightarrow{k_2} C$.

Typical concentration-time patterns for A, B and C are shown in Fig. 1.13. The differential equations and mass balance that describe this reaction are:

$$\frac{d[A]}{dt} = -k_1[A] \tag{1.67}$$

$$\frac{d[B]}{dt} = k_1[A] - k_2[B] \tag{1.68}$$

$$\frac{d[C]}{dt} = k_2[B] \tag{1.69}$$

$$[A_o] + [B_o] + [C_o] = [A_t] + [B_t] + [C_t] \tag{1.70}$$

Once the differential equations and mass balance have been written down, three approaches can be followed in order to model complex reaction schemes. These are:

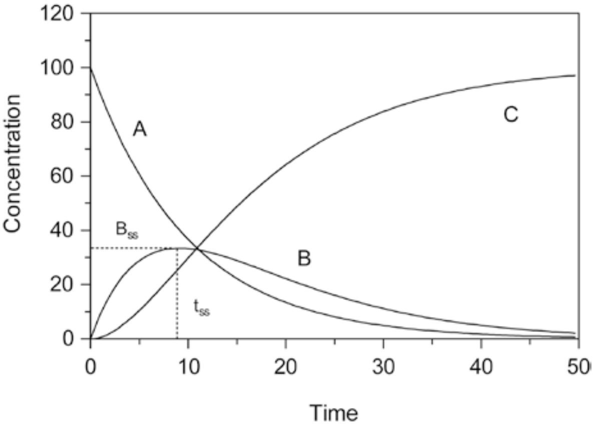


Fig. 1.13 Changes in reactant, intermediate, and product concentrations as a function of time for a reaction of the form $A \to B \to C$. B_{ss} denotes the steady-state concentration in intermediate B at time t_{ss}

1. steady-state approximations to solve differential equations analytically
2. exact analytical solutions of the differential equations without using approximations
3. numerical integration of differential equations

It is important to remember that in this day and age of powerful computers, it is not necessary to find analytical solutions to differential equations anymore. Many commercially available software packages will carry out numerical integration of differential equations followed by nonlinear regression to fit the model, in the form of differential equations, to the data. Estimates of the rate constants and their variability, as well as measures of the goodness of fit of the model to the data, can be obtained in this fashion. Eventually all modeling exercises are carried out in this fashion since it is difficult, and sometimes impossible, to obtain analytical solutions for complex reaction schemes.

1.8.1 Exact Analytical Solution (Non-steady State Approximation)

Exact analytical solutions for the reaction $A \rightarrow B \rightarrow C$ can be obtained by solving the differential equations using standard mathematical procedures or by merely referring to the appropriate page of a common mathematical handbook. Exact solutions to the differential equations for the boundary conditions $[B_o] = [C_o] = 0$ at $t = 0$, and therefore $[A_o] = [A_t] + [B_t] + [C_t]$, are:

$$[A_t] = [A_o] e^{-k_1 t} \tag{1.71}$$

$$[B_t] = k_1 [A_o] \left(\frac{e^{-k_1 t} - e^{-k_2 t}}{k_2 - k_1} \right) \tag{1.72}$$

$$[C_t] = [A_o] \left[1 + \frac{1}{k_1 - k_2} \left(k_2 e^{-k_1 t} - k_1 e^{-k_2 t} \right) \right] \tag{1.73}$$

Figure 1.13 shows the simulation of concentration changes in the system $A \rightarrow B \rightarrow C$. The models (equations) are fitted to the experimental data using nonlinear regression in order to obtain estimates of k_1 and k_2.

1.8.2 Exact Analytical Solution (Steady State Approximation)

Steady-state approximations are useful, and thus extensively used in the development of mathematical models of kinetic processes. Take, for example, the reaction $A \rightarrow B \rightarrow C$ (Fig. 1.13). If the rate at which A is converted to B equals the rate at which B is converted to C, then the concentration of B remains constant, or in a steady-state. It is important to remember that molecules of B are constantly being created and destroyed, but since these processes are occurring at the same rate, the net effect is that the concentration of B remains unchanged $\left(\frac{d[B]}{dt}=0\right)$, thus:

$$\frac{d[B]}{dt}=0=k_1[A]-k_2[B] \tag{1.74}$$

Decreases in $[A]$ as a function of time are modeled as a first-order decay process,

$$[A_t]=[A_o]e^{-k_1 t} \tag{1.75}$$

with the value of k_1 being determined as previously discussed.

From Eqs. 1.79 and 1.80, we can deduce that:

$$[B]=\frac{k_1}{k_2}[A]=\frac{k_1}{k_2}[A_o]e^{-k_1 t} \tag{1.76}$$

If the steady-state concentration of B ($[B]_{ss}$), the value of k_1, and the time at which that steady-state was reached (t_{ss}) are known, k_2 can be determined from Eq. 1.82:

$$k_2=\frac{k_1}{[B]_{ss}}[A_o]e^{-k_1 t_{ss}} \tag{1.77}$$

The steady-state of B in the reaction $A \rightarrow B \rightarrow C$ is short lived. However, for many reactions, such as enzyme-catalyzed reactions, the concentrations of important reaction intermediates are in a steady-state. This allows for the use of steady-state approximations to obtain analytical solutions for the differential equations and thus be able to estimate the values of the rate constants.

1.8.3 Numerical Integration and Regression

1.8.3.1 Numerical Integration

Finding the numerical solution of a system of a first order ordinary differential equations,

$$\frac{dY}{dx}=f\left(x,Y(x)\right)\text{where } Y(x_o)=Y \tag{1.78}$$

entails finding the numerical approximations of the solution $Y(x)$ at discrete points $x_o, x_1, x_2 < \ldots < x_n < x_{n+1} < \ldots$ by $Y_o, Y_1, Y_2, \ldots, Y_n, Y_{n+1}, \ldots$ The distance between two consecutive points, $h_n = x_n - x_{n+1}$, is called the step size. Step sizes do not necessarily have to be constant between all grid points x_n. All numerical methods have one property in common, namely, finding approximations of the solution $Y(x)$ at grid points one by one. Thus, if a formula can be given to calculate Y_{n+1} based on the information provided by the known values of $Y_n, Y_{n-1}, \ldots, Y_o$, then the problem is solved. Many numerical methods have been developed to find solutions for ordinary differential equations, the simplest one being the Euler method. Even though the Euler method is seldom used in practice due to lack of accuracy, it serves as the basis for analysis in more accurate methods, such as the Runge-Kutta method, among many others.

For a small change in the dependent variable (Y) in time (x), the following approximation is used:

$$\frac{dY}{dx} \approx \frac{\Delta Y}{\Delta x} \tag{1.79}$$

Therefore we can write:

$$\frac{\left(Y_{n+1} - Y_n\right)}{\left(x_{n+1} - x_n\right)} = f\left(x_n, Y_n\right) \tag{1.80}$$

By rearranging the above equation, Euler obtained an expression for Y_{n+1} in terms of Y_n:

$$Y_{n+1} = Y_n + \left(x_{n+1} - x_n\right) f\left(x_n, Y_n\right) \text{ or } Y_{n+1} = Y_n + h f\left(x_n, Y_n\right) \tag{1.81}$$

Consider the reaction $A \rightarrow B \rightarrow C$. As discussed above, the analytical solution for the differential equation which describes the first order decay in $[A]$ is $[A_t] = [A_o]e^{-kt}$ and thus the differential equation that describes changes in $[B]$ in time can be written as:

$$\frac{d[B]}{dt} = k_1 [A_o] e^{-k_r t} - k_2 [B] \tag{1.82}$$

A numerical solution for the above differential equation is found using the initial value $[B_o]$ at $t = 0$, and from the knowledge of the values of k_1, k_2, and $[A_o]$. Values for $[B_t]$ are then calculated as follows:

$$\frac{dB}{dt} = \frac{\Delta B}{\Delta t} = k_1 [A_o] - k_2 [B] \tag{1.83}$$

$$\Delta B = \Delta t \left(k_1 [A_o] - k_2 [B]\right) \tag{1.84}$$

$$\Delta B = B_1 - B_o \tag{1.85}$$

$$B_1 - B_o = \Delta t \left(k_1 [A_o] - k_2 [B]\right) \tag{1.86}$$

$$[B_1] = [B_o] + h\left(k_1 [A_o] - k_2 [B_o]\right) \tag{1.87}$$

$$[B_2]=[B_1]+h\left(k_1[A_o]e^{-k_1t_1}-k_2[B_1]\right) \tag{1.88}$$

$$\vdots$$

$$[B_{n+1}]=[B_n]+h\left(k_1[A_o]e^{-k_1t_n}-k_2[B_n]\right) \tag{1.89}$$

It is therefore possible to generate a numerical solution (i.e. a set of numbers predicted by the differential equation) of the above ordinary differential equation. Values obtained from the numerical integration, i.e. predicted data, can now be compared to experimental data values.

1.8.3.2 Least-Squares Minimization (Regression Analysis)

The most common way in which models are fitted to data is by using least-squares minimization procedures (regression analysis). All these procedures, linear or nonlinear, seek to find estimates of the equation parameters $(\alpha,\beta,\gamma,\delta,\ldots)$ by determining parameter values for which the sum of squared residuals is at a minimum, and therefore:

$$\left(\frac{\partial\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}{\partial\alpha}\right)_{\beta,\gamma,\delta,\ldots}=0 \tag{1.90}$$

where y_i and $\hat{y}_i$ correspond, respectively, to the ith experimental and predicted points at x_i. If the variance (s_i^2) of each data point is known from experimental replication, then a ‘weighted’ least-squares minimization can be carried out, where the weights (w_i) correspond to $1/s_i^2$. In this fashion, data points that have greater error contribute less to the analysis, namely

$$\left(\frac{\partial\sum_{1}^{n}w_i\left(y_i-\hat{y}_i\right)^2}{\partial\alpha}\right)_{\beta,\gamma,\delta,\ldots.}=0 \tag{1.91}$$

Regression analysis involves several important assumptions about the function chosen and the error structure of the data. These include:

(a) *The correct equation is used.*
(b) *Only dependent variables are subject to error, while independent variables are known exactly.*

(c) *Errors are normally distributed with zero mean, are the same for all responses (homoscedastic errors), and are uncorrelated (zero covariance).*
(d) *The correct weighting is used.*

For linear functions, single or multiple, it is possible to find analytical solutions of the error minimization partial differential. Therefore, exact mathematical expressions exist for the calculation of a slopes and intercepts. It should be noted at this point that a linear function of parameters does not imply a straight line. A model is linear if the first partial derivative of the function with respect to the parameter(s) is independent of such parameter(s), therefore, higher order derivatives would be zero.

For example, equations used to calculate the 'best-fit' slope and y-intercept for a data set that fits the linear function $y = mx+b$ can be easily obtained by considering that the minimum sum of squared residuals (SS) corresponds to parameter values for which the partial differential of the function with respect to each parameter equals zero. The squared residuals to be minimized are:

$$(\text{Residual})^2 = \left(y_i - \hat{y}_i\right)^2 = \left[y_i - \left(mx_i + b\right)\right]^2 \tag{1.92}$$

The partial differential of the slope (m), for a constant y-intercept (b) is therefore:

$$\left(\frac{\partial SS}{\partial m}\right)_b = -2\sum_1^n x_i y_i + 2b\sum_1^n x_i + 2m\sum_1^n x_i^2 = 0 \tag{1.93}$$

And thus,

$$m = \frac{\sum_1^n x_i y_i - b\sum_1^n x_i}{\sum_1^n x_i^2} \tag{1.94}$$

The partial differential of the y-intercept (b), for a constant slope (m) is:

$$\left(\frac{\partial SS}{\partial b}\right)_m = m\sum_1^n x_i - \sum_1^n y_i + nb = 0 \tag{1.95}$$

And therefore,

$$b = \frac{\sum_1^n y_i - m\sum_1^n x_i}{n} = \overline{y} - m\overline{x} \tag{1.96}$$

where $\overline{x}$ and $\overline{y}$ correspond to the overall averages of all x and y data, respectively. Substituting b into m, and rearranging, we obtain an equation for the direct calculation of the 'best fit' slope of the line:

$$m = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}}{\sum_{i=1}^{n} x_i^{\,2} - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{1.97}$$

The ‘best-fit’ y-intercept of the line is given by:

$$b = \bar{y} - \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \bar{x} \tag{1.98}$$

At this point, it would be useful to mention goodness-of-fit statistics. A useful parameter for judging the ‘goodness of fit’ of a model to experimental data is the reduced chi square (χ_v^2) statistic:

$$\chi_v^2 = \frac{\sum_{1}^{n} w_i \left(y_i - \hat{y}_i \right)^2 / \hat{y}_i}{v} \tag{1.99}$$

where w_i is the weight of the i^{th} data point ($w_i = 1/s_i^2$, where s_i^2 is the variance of mean of y_i values, and ν corresponds to the degrees of freedom, defined as $\nu=(n-p-1)$, where n is the total number of data values and p is the number of parameters that are estimated. The reduced χ_v^2 value should be roughly equal to the number of degrees of freedom if the model is correct, i.e., $\chi_v^2 \approx 1$.

Another statistic most appropriately applied to linear regression, as an indication of how closely the dependent and independent variables approximate a linear relationship to each other, is the correlation coefficient (CC):

$$CC = \frac{\sum_{i=1}^{n} w_i (x_i - \bar{x})(y_i - \bar{y})}{\left(\sum_{i=1}^{n} w_i (x_i - \bar{x})^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} w_i (y_i - \bar{y})^2\right)^{\frac{1}{2}}} \tag{1.100}$$

Values for the correlation coefficient can range from −1 to +1. A CC value close to +/−1 is indicative of a strong correlation.

The coefficient of determination (CD) is the fraction of the total variability accounted for by the model. This is a more appropriate measure of the goodness-of-fit of a model to data than the r^2 statistic. The CD has the general form:

$$CD = \frac{\sum_{i=1}^{n} w_i (y_i - \bar{y})^2 - \sum_{i=1}^{n} w_i \left(y_i - \hat{y}_i \right)^2}{\sum_{i=1}^{n} w_i (y_i - \bar{y})^2} \tag{1.101}$$

Finally, the r^2 statistic is similar to the CD. This statistic is often used erroneously when, strictly speaking, the CD should be used. The root of the r^2 statistic is sometimes erroneously reported to correspond to the CD. An r^2 value close to +/−1 is indicative that the model accounts for most of the variability in the data. The r^2 statistic has the general form:

$$r^2 = \frac{\sum_{i=1}^{n} w_i y_i^2 - \sum_{i=1}^{n} w_i \left(y_i - \hat{y}_i \right)^2}{\sum_{i=1}^{n} w_i y_i^2} \tag{1.102}$$

1.8.3.3 Non-linear Regression: Techniques and Philosophy

For nonlinear functions, however, the situation is more complex. Iterative methods are used instead, in which parameter values are changed simultaneously, or one at a time, in a prescribed fashion until a global minimum is found. The algorithms used include the Levenberg-Marquardt method, the Powell method, the Gauss-Newton method, the steepest-descent method, simplex minimization, and combinations thereof. It is beyond the scope of this chapter to discuss the intricacies of procedures used in nonlinear regression analysis. Suffice to say, most modern graphical software packages include nonlinear regression as a tool for curve-fitting.

At this point it is necessary to discuss differences between uniresponse and multiresponse modelling. Take for example the reaction A→B→C. Usually, equations in differential or algebraic form, are fitted to individual data sets, A, B, and C, and a set of parameter estimates obtained. However, if changes in the concentrations of A, B and C as a function of time are determined, it is possible to use the entire data set (A, B, C) simultaneously to obtain parameter estimates. This procedure entails fitting the functions that describe changes in the concentration of A, B and C to the experimental data *simultaneously*, thus obtaining one global estimate of the rate constants. This, so called, multivariate response modeling helps increase the precision of the parameter estimates by using all available information from the various responses. However, this type of analysis is beyond the scope of this book. It is not difficulty to carry out, but it does require experience and perspective.

Should the reader decide to embark on a modelling journey, some comments on curve-fitting and nonlinear regression are required. There is no general method that guarantees obtaining the best global solution to a nonlinear least squares minimization problem. Even for a single parameter model, several minima may exist! A minimization algorithm will eventually succeed in find *a minimum*, however, there is no assurance that this corresponds to the *global minimum*. It is theoretically possible for one, and maybe two parameter functions, to exhaustively search all parameter initial values and find the global minimum. However, this approach is usually not practical even beyond a single parameter function.

There are, however, some guidelines that can be followed to increase the likelihood of finding the best fit to nonlinear models. All nonlinear regression algorithms require *initial estimates* of parameter values. These initial estimates should be as close as possible to their best-fit value so that the program can actually succeed in finding the global minimum. The development of good initial estimates comes primarily from the scientists' physical knowledge of the problem at hand, as well as from intuition and experience. Curve-fitting can sometimes be somewhat of an art form.

Generally, it is useful to carry out simulations varying initial estimates of parameter values in order to develop a 'feel' of how changes in initial estimate values will affect the nonlinear regression results obtained. Some programs offer simplex minimization algorithms that do not require the input of initial estimates. These secondary minimization procedures may provide values of initial estimates for the primary minimization procedures. Once a minimum is found, there is no assurance, however, that it corresponds to the global minimum. A standard procedure to test whether the global minimum has been reached is called *sensitivity analysis*. Sensitivity analysis refers to the variability in results (parameter estimates) obtained from nonlinear regression analysis due to changes in the values of initial estimates. In sensitivity analysis, least-squares minimizations are carried out for different starting values of initial parameter estimates in order to determine whether the convergence to the same solution is attained. If the same minimum is found for different values of initial estimates, then the scientist can be fairly confident that the proposed minimum is the best answer. Another approach is to fit the model to the data using different weighting schemes, since it is possible that the largest or smallest values in the data set may have an undue influence on the final result. Very important as well is the visual inspection of the data and plotted curve(s), since a graph can provide clues that may aid in finding a better solution to the problem.

Strategies exist for systematically finding minima and hence finding the best minimum. In a multi-parameter model, it is sometimes useful to vary one or two parameters at a time. This entails carrying out the least-squares minimization procedure floating one parameter at a time while fixing the value of the other parameters as constants, and/or analyzing a subset of the data. This simplifies calculations enormously, since the greater the number of parameters to be simultaneously estimated, the more difficult it will be for the program to find the global minimum. For example, for the reaction $A \rightarrow B \rightarrow C$, k_1 can be easily estimated from the first order decay of [A] in time. The parameter k_1 can therefore be fixed as a constant, and only k_2 and k_3 floated. After preliminary parameter estimates are obtained in this fashion, these parameters should be fixed as constants and the remaining parameters estimated. Only after estimates are obtained for all the parameters should the entire parameter set be fitted simultaneously. It is also possible to assign physical limits, or constraints, to the values of the parameters. The program will find a minimum that corresponds to parameter values within the permissible range.

Care should be exercised at the data-gathering stage as well. A common mistake is to gather all the experimental data without giving much thought as to how the data will be analyzed. It is extremely useful to use the model to simulate data sets, and then try to fit the model to the simulated data. This exercise will promptly point out where more data would be useful to the model building process. It is a good investment of time to simulate the experiment and data analysis to identify where problems may lie, and identify regions of data that may be most important in determining the properties of the model. The gathered data must be amenable to analysis in such a way as to shed light on the model.

For difficult problems, the determination of best-fit parameters is a procedure that benefits greatly from experience, intuition, perseverance, skepticism and scientific reasoning. A good answer requires good initial estimates. Start the minimization procedure with the best possible initial estimates for parameters, and if the parameters have physical limits, specify constraints on their value. For complicated models, begin model fitting by floating a single parameter and using a subset of the data that may be most sensitive to changes in the value of the particular parameter. Subsequently add parameters and data until it is possible to fit the full model to the complete data set. After the minimization is accomplished, test the answers by carrying out sensitivity analysis. Maybe run a simplex minimization procedure to determine if there are other minima nearby, and whether or not the minimization 'wanders off' in another direction. Finally plot the data and calculated values and visually check for goodness of fit – the human eye is a powerful tool. Above all, care should be exercised – if curve-fitting is approached blindly without understanding its inherent limitations and nuances, erroneous results will be obtained.

The F-test is the most common statistical tool used to judge whether a model fits the data better than another. The models to be compared are fitted to data and reduced χ_v^2 values obtained. The ratio of the χ_v^2 values obtained is the F-statistic, namely:

$$F_{v_a,v_b} = \frac{\chi_v^2(a)}{\chi_v^2(b)} \tag{1.103}$$

where ν stands for the degrees of freedom, which are determined from $\nu = n - p - 1$, where n and p correspond, respectively, to the total number of data points and the number of parameters in the model. Using standard statistical tables, it is possible to determine if the fits of the models to the data are significantly different from each other at a certain level of statistical significance.

The analysis of residuals $(\hat{y}_i\text{-}y_i)$, in the form of the serial correlation coefficient, provides a useful measure of how much the model deviates from the experimental data. Serial correlation is an indication of whether residuals tend to run in groups of positive or negative values, or tend to be scattered randomly about zero. A large positive value of the serial correlation coefficient is indicative of a systematic deviation of the model from the data. The serial correlation coefficient (SCC) has the general form:

$$SCC = \sqrt{n-1}\,\frac{\sum_{i=1}^{n}\sqrt{w_i}\left(\hat{y}_i - y_i\right)\sqrt{w_{i-1}}\left(\hat{y}_{i-1} - y_{i-1}\right)}{\sum_{i=1}^{n}\left(w_i\left(\hat{y}_i - y_i\right)\right)^2} \tag{1.104}$$

A word of caution about weighting of data needs to be mentioned. Consider a typical experiment where the value of a dependent variable is measured several times at a particular value of the independent variable. If the experiment itself is then replicated several times, a set of sample means and variances of *sample means* can be obtained. As stated before, the weight is merely the inverse of this sample variance. The two most basic assumptions made in regression analysis are that experimental errors are normally distributed with mean zero, and that errors are the same for all data points (error homoscedasticity). Systematic trends in the experimental errors, or the presence of outliers would invalidate these assumptions. The purpose of weighting residuals is to eliminate systematic trends in the error structure of the data, and "weigh less" excessively noisy data. Thus, the error structure present in the experimental data needs to be determined, which is not a trivial task. Ideally, each experiment would be replicated sufficiently so that individual data weights could be calculated directly from experimentally determined variances. However, replicating experiments to the extent that would be required in order to obtain accurate estimates of the errors is expensive, time consuming, and impractical. It is important to note that if insufficient data points are used to estimate individual errors of data points, incorrect estimates of weights will be obtained. The use of incorrect weights in regression analysis will make matters worse – if in doubt, do not weigh the data.

1.9 Enzyme Kinetics

1.9.1 Enzyme Catalyzed Reactions

An enzyme is a protein with catalytic properties. As a catalyst, an enzyme lowers the energy of activation for a reaction (E_a). This increases the rate of reaction without affecting the position of equilibrium. An enzyme increases the rate of a reaction primarily by specifically binding to, and thus stabilizing, the transition state structure. Joseph Kraut eloquently pointed this out when he stated that "an enzyme can be considered a flexible molecular template, designed by evolution to be precisely complementary to the reactants in their activated transition-state geometry, as distinct from their ground-state geometry. Thus an enzyme strongly binds the transition state, greatly increasing its concentration, and accelerating the reaction proportionately." This description of enzyme catalysis is now usually referred to as *transition-state stabilization*.

1.9.2 Characterizing Enzyme Activity

The *enzyme unit* (e.u.) is the most commonly used standard unit of enzyme activity. An enzyme unit is defined as the amount of enzyme that causes the disappearance of 1 μ mol of substrate per minute. As well, it can be defined in terms of the moles of product that appears per minute, $1\,\text{e.u.} = \frac{1\,\mu mol}{\text{min}}$

Specific activity is defined as the number of enzyme units per unit mass. This mass could correspond to the mass of the pure enzyme, the amount of protein in a particular isolate, or the total mass of the tissue from where the enzyme was derived. Regardless of which case it is, this must be clearly stated.

The *molecular activity* or the *turnover number*, k_{cat}, is the number of substrate molecules converted to product per enzyme molecule per unit time when the enzyme is fully saturated with substrate.

In order to determine enzyme-catalyzed reaction velocities, it is necessary to generate a progress curve. For the conversion of the substrate S to the product P, the general shape of the progress curve is that of a first-order exponential decrease in substrate concentration (Fig. 1.14):

$$[S - S_{\min}] = [S_o - S_{\min}]e^{-kt} \tag{1.105}$$

or that of a first-order exponential increase in product concentration (Fig. 1.14):

$$[P - P_o] = [P_{\max} - P_o]\left(1 - e^{-kt}\right) \tag{1.106}$$

where $[S_o]$, $[S_{min}]$ and $[S]$ correspond, respectively, to the initial substrate concentration at $t=0$, the minimum substrate concentration as $t \to \infty$, and the substrate concentration at time t, while $[P_o]$, $[P_{max}]$ and $[P]$ correspond to, respectively, the initial product concentration at $t=0$, the maximum product concentration as $t \to \infty$, and the product concentration at time t.

The rate of the reaction, or reaction velocity (v), corresponds to the instantaneous slope of either of the progress curves, namely:

$$v = -\frac{dS}{dt} = \frac{dP}{dt} \tag{1.107}$$

However, as can be appreciated in Fig. 1.14, the reaction velocity, i.e., the slope of the curve, decreases in time. This drop can be caused by any number of the following reasons:

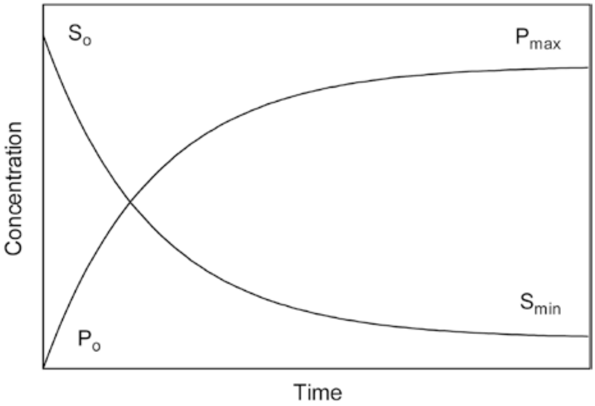


Fig. 1.14 Changes in substrate (S) and product (P) concentration as a function of time, from initial values (S_0 and P_0) to final values (P_{max} and S_{min})

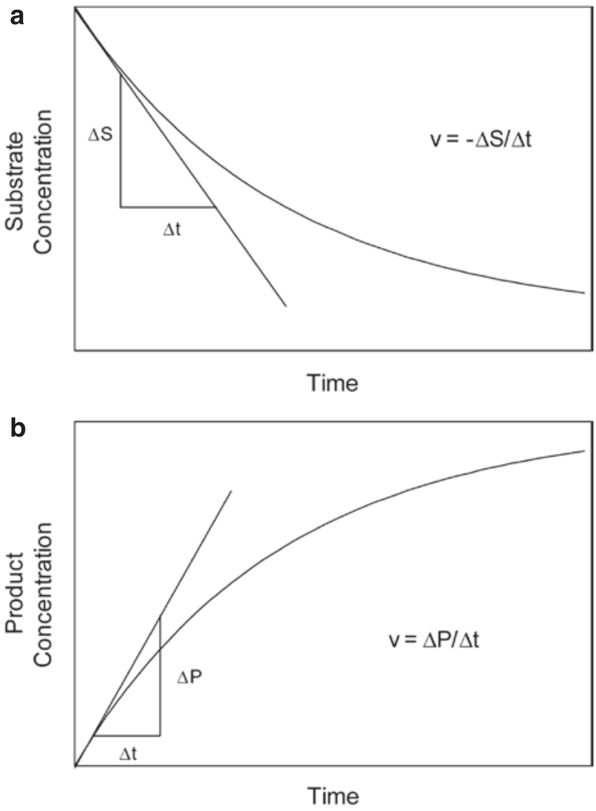


Fig. 1.15 Determination of the initial velocity of an enzyme-catalyzed reaction from the instantaneous slope at t = 0 of substrate depletion (**a**) or product accumulation (**b**) progress curves

1. The enzyme becomes unstable during the course of the reaction.
2. The degree of saturation of the enzyme by substrate decreases as substrate is depleted.
3. The reverse reaction becomes predominant once product accumulates.
4. The reaction products inhibit the enzyme (negative feedback).

It is for these reasons that progress curves for enzyme-catalyzed reactions do not fit standard models for homogeneous chemical reactions requiring a different approach. Enzymologists use initial velocities as a measure of reaction rates instead. During the early stages of an enzyme-catalyzed reaction, conversion of substrate to product is small, and can thus be considered to remain constant and effectively equal to the initial substrate concentration ($[S_t] \approx [S_o]$). By the same token, very little product has accumulated ($[P_t] \approx 0$); thus, the reverse reaction can be considered to be negligible, and any possible inhibitory effects of product on enzyme activity, insignificant. More importantly, the enzyme can be considered to remain stable during the early stages of the reaction. In order to obtain initial velocities, a tangent to the progress curve is drawn as close as possible to its origin (Fig. 1.15). The slope of this tangent, i.e., the initial velocity, is obtained using linear regression.

For proper enzyme kinetic analysis, it is essential to obtain reaction velocities strictly from the initial region of the progress curve. By using the wrong time for the

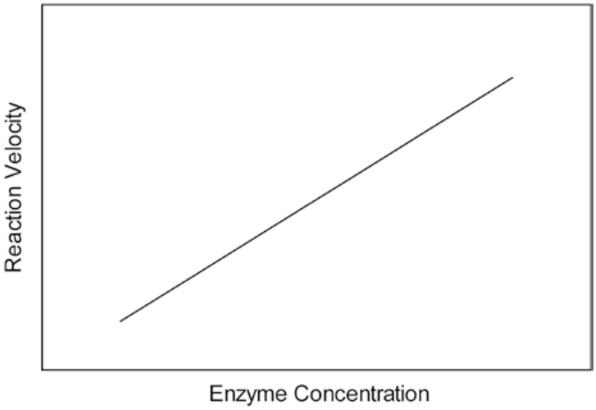


Fig. 1.16 Linear dependence of reaction initial velocity on enzyme concentration in the reaction mixture.

derivation of rates (not necessarily initial velocities), a linear relationship between enzyme concentration and velocity will not be obtained, this being a basic requirement for enzyme kinetic analysis. For the reaction to be kinetically controlled by the enzyme, reaction velocity must be directly proportional to enzyme concentration (Fig. 1.16).

Continuous and discontinuous methods used to monitor the progress of an enzymatic reaction may not always agree. This can be the case particularly for two-stage reactions in which an intermediate between product and substrate accumulates. In this case, the disappearance of substrate may be a more reliable indicator of activity than product accumulation. For discontinuous methods, at least three points are required, one at the beginning of the reaction ($t = 0$), one at a convenient time t_1, and one at time t_2, which should correspond to twice the length of t_1. This three-point measurement provides a check for linearity.

1.9.3 The Equilibrium Catalysis Model

An enzymatic reaction is usually modeled as a two-step process – substrate (S) binding by enzyme (E) and formation of an enzyme-substrate (ES) complex, followed by an irreversible breakdown of the enzyme-substrate complex to free enzyme and product (P), namely:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P \tag{1.108}$$

In the equilibrium model of Michaelis and Menten, the substrate-binding step is assumed to be fast relative to the rate of breakdown of the ES complex. Therefore, the substrate binding reaction is assumed to be at equilibrium. The equilibrium dissociation constant for the ES complex, (K_s), is a measure of the affinity of enzyme for substrate. The lower the value of K_s, the greater the affinity of the enzyme for the substrate. The K_s corresponds to substrate concentration at $\frac{1}{2}V_{\max}$:

$$K_s = \frac{[E][S]}{[ES]} \tag{1.109}$$

The rate limiting step for the particular series of reactions is the breakdown of the ES complex, which can be expressed in the following rate equation:

$$v = \frac{d[ES]}{dt} = k_{cat}[ES] \tag{1.110}$$

where k_{cat} corresponds to the effective first-order rate constant for the breakdown of the ES complex to free product and free enzyme. The rate equation is usually normalized by the total enzyme concentration ($[E_T] = [E] + [ES]$) to give:

$$\frac{v}{[E_T]} = \frac{k_{cat}[ES]}{[E]+[ES]} \tag{1.111}$$

where [E] and [ES] correspond, respectively, to the concentrations of free enzyme and enzyme-substrate complex. By re-arranging Eq. 1.111 and substituting $\frac{[E][S]}{K_s}$ into [ES] yields:

$$\frac{v}{[E_T]} = \frac{k_{cat}\frac{[E][S]}{K_s}}{[E]+\frac{[E][S]}{K_s}} \tag{1.112}$$

After dividing both the numerator and denominator by [E], multiplying the numerator and denominator by K_s and then re-arranging, the familiar Michaelis-Menten expression for the velocity of an enzyme catalyzed reaction is obtained:

$$v = \frac{k_{cat}[E_T][S]}{K_s + [S]} \tag{1.113}$$

$k_{cat}[E_T]$ can be redefined as the maximum reaction velocity V_{max}. Eq. 1.113 can be re-written as:

$$v = \frac{V_{\max}[S]}{K_s + [S]} \tag{1.114}$$

The Michaelis-Menten model makes the following assumptions:

1. The rate of the formation of the ES complex is fast relative to its breakdown rate.
2. The concentration of substrate remains constant during the time period in which the velocity is measured.
3. The conversion of product to substrate is negligible.

In cases where rapid equilibrium conditions cannot be assumed, a steady-state approximation can be used to obtain an enzyme catalysis model.

1.9.4 The Steady-State Catalysis Model

The main assumption made in the steady-state approximation is that the concentration of enzyme-substrate complex remains constant in time, i.e., $\frac{d[ES]}{dt} = 0$. Thus, the differential equation that describes changes in the concentration of the ES complex in time equals zero:

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_2[ES] = 0 \tag{1.115}$$

Re-arranging this equation gives an expression for K_m, the Michaelis constant:

$$K_m = \frac{[E][S]}{[ES]} = \frac{k_{-1} + k_2}{k_1} \tag{1.116}$$

This K_m will only be equivalent to the dissociation constant of the ES complex (K_s), used in the equilibrium catalysis model, for the case where $k_{-1} >> k_2$, and therefore $K_m = \frac{k_{-1}}{k_1} = K_s$. The Michaelis constant K_m corresponds to substrate concentration at $\frac{1}{2} V_{\max}$.

As stated before, the rate-limiting step of an enzyme-catalyzed reaction is the breakdown of the ES complex. The velocity of the enzymatic reaction can thus be expressed as:

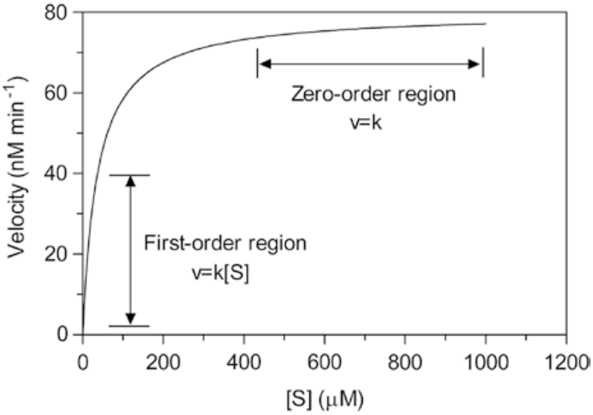
$$v = k_{cat}[ES] \tag{1.117}$$

The rest of the derivation is the same as for the equilibrium model.

1.9.5 The Initial Velocity vs [S] Plot

The general shape of a velocity versus substrate concentration curve is that of a rectangular hyperbola (Fig. 1.17). At low substrate concentrations, the rate of the reaction is proportional to the substrate concentration. In this region, the order of the enzymatic reaction is first order with respect to substrate concentration (Fig. 1.17). For the case where $[S] = K_m$, the Michaelis-Menten model will reduce to:

Fig. 1.17 Initial velocity versus substrate concentration plot for an enzyme-catalyzed reaction. Notice the first- and zero-order regions of the curve, where the reaction velocity is, respectively, linearly dependent and independent of substrate concentration



$$v = \frac{k_{cat}}{K_m}[E_T][S] = \frac{V_{\max}}{K_m}[S] \tag{1.118}$$

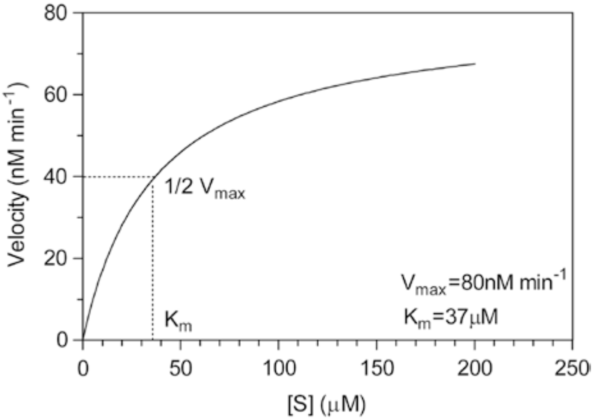
where $\frac{k_{cat}}{K_m}$ [$M^{-1}s^{-1}$] is the second order rate constant for the reaction, while $\frac{V_{\max}}{K_m}$ [s^{-1}] is the first order rate constant for the reaction. Knowledge of enzyme concentration allows for the calculation of $\frac{k_{cat}}{K_m}$ from $\frac{V_{\max}}{K_m}$. There are some physical limits to this ratio. The ultimate limit on the value of $\frac{k_{cat}}{K_m}$ is dictated by k_1. This step is solely controlled by the rate of diffusion of substrate to the active site of the enzyme. This, in turn is related to the solvent viscosity. This limits the value of k_1 to 10^8 to 10^9 $M^{-1}s^{-1}$. The ratio $\frac{k_{cat}}{K_m}$ for many enzymes is in this range. This suggests that the catalytic activity of many enzymes depends solely on the rate of diffusion of the substrate to the active site.

At higher concentrations, the velocity of the reaction remains approximately constant, and effectively insensitive to changes in substrate concentration. At this point, all enzyme molecules are saturated with substrate. In this region, the order of the enzymatic reactions is zero order with respect to substrate (Fig. 1.17). For the case where $[S] >> K_m$, Eq. 1.113 will reduce to:

$$v = k_{cat}[E_T] = V_{\max} \tag{1.119}$$

The value of K_m varies widely. For most enzymes, it generally lies between 10^{-1} M and 10^{-7} M. The value of K_m depends on the type of substrate and on environmental conditions such as pH, temperature, ionic strength and polarity. K_m and K_s correspond to the concentration of substrate at half maximum velocity (Fig. 1.18). This fact can be readily shown by substituting [S] with K_s in Eq. 1.119. It is important to remember that K_m equals K_s only when the breakdown of the ES complex takes

Fig. 1.18 Initial velocity versus substrate concentration plot for an enzyme with V_{max} = 80 nM/min and K_m = 37 μM



place much more slowly than the binding of substrate to the enzyme; that is, when $k_{-1} >> k_2$, and thus:

$$K_m = \frac{k_{-1}}{k_1} = K_s \tag{1.120}$$

1.9.6 Determining Parameters of an Enzyme Kinetic Model: Method 1

The first step to determining parameters for the Michaelis-Menten model of an enzyme is to validate whether the assumptions for the Michaelis-Menten model have been met, specifically:

1. The activity of the enzyme does not decrease over the time course of the experiment, i.e. the enzyme remains stable.
2. Reaction velocity should be linearly proportional to the enzyme concentration.

Enzyme stability can be tested by using Selwyn's simple assay for enzyme inactivation. In Selwyn's test, the reaction's progress (in %) is plotted as a function of the product of the time t and the total enzyme concentration $[E_o]$, $[E_o]t$. If plots with different enzyme concentrations are superimposable, this indicates that the enzyme is not being inactivated over the course of the experiment. Non-superimposable graphs imply that the reaction rate is not linearly proprotional to the enzyme concentration. This, in turn, indicates that the enzyme is being inactivated.

In the experimental setup, substrate concentrations in the range $0.5K_m$–$10K_m$ should, if possible, be used. These should be spaced more closely at low substrate concentrations, with at least one high concentration approaching V_{max}. Concentrations of $\frac{1}{3}$, $\frac{1}{2}$, 1, 2, 4, 8 K_m are appropriate. Each substrate concentration must preferably have three replicates.

As will be later discussed, the collected velocity and substrate concentration data can then be fitted to the Michaelis-Menten model in order to obtain estimates of K_m and V_{max}.

1.9.7 Determining Parameters of an Enzyme Kinetic Model: Method 2

It is theoretically possible to derive estimates of K_m and V_{max} from a single progress curve. This method is preferable to that described in the previous chapter since it does not require measuring the initial velocity as a function of several different concentrations, hence minimizing the number of experiments that must be performed. What follows is the explicit solution of the Michealis-Menten differential equation.

The velocity of an enzyme-catalyzed reaction as given by the Michaelis-Menten model can be described in terms of the disappearance of substrate:

$$-\frac{d[S]}{dt}=\frac{V_{\max}[S]}{K_m+[S]} \tag{1.121}$$

By multiplying the given differential equation by $\frac{(K_m+[S])}{[S]}dt$ and subsequent integration over the boundary conditions $[S]=[S_o]$ at $t=0$ and $[S]=[S_t]$ at time t gives:

$$-K_m\int_{S_o}^{S_t}\frac{d[S]}{[S]}-\int_{S_o}^{S_t}d[S]=V_{\max}\int_0^t dt \tag{1.122}$$

After performing the necessary integration operations, the following expression is obtained:

$$K_m\ln\frac{S_o}{S_t}+[S_o-S_t]=V_{\max}t \tag{1.123}$$

The use of commercial software to regress the above equation to a given data set often requires that $[S]$ be an explicit function of t. Algebraic manipulation of the equation gives:

$$\frac{1}{t}\ln\frac{[S_0]}{[S_t]}=-\frac{[S_0-S_t]}{K_m t}+\frac{V_{\max}}{K_m} \tag{1.124}$$

A plot of $\frac{1}{t}\ln\frac{S_o}{S_t}$ versus $\frac{[S_o-S_t]}{t}$ yields a straight line with a slope of $-\frac{1}{K_m}$, x-intercept = V_{max} and y-intercept = $\frac{V_{\max}}{K_m}$ as shown in Fig. 1.19. The values of the

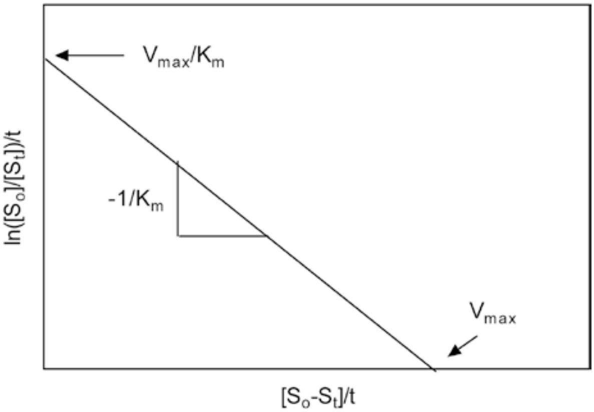


Fig. 1.19 A linear plot of a single progress curve used to determine the catalytic parameters V_{max} and K_m.

slope and intercept can be readily obtained using linear regression making it possible to obtain the required kinetic parameters from a single $[S_t]$-t data set.

1.9.8 *Kinetic Effects of Reversible Inhibition*

An *inhibitor* is a compound that decreases the rate of an enzyme-catalyzed reaction. This inhibition can be reversible or irreversible. Reversible enzyme inhibition can be competitive, uncompetitive or non-competitive, each affecting the K_s and V_{max} in a specific fashion. In this chapter, each type of reversible inhibition will be discussed in turn. This will be followed by a discussion of the methodology used to measure the enzyme-inhibitor dissociation constant (K_i).

1.9.8.1 Competitive Inhibition

A *competitive inhibitor* competes with the enzyme's substrate for binding to the active site. The mechanism is illustrated in Fig. 1.20.

The result is an increase in the enzyme-substrate dissociation constant K_s which signifies a decrease in the affinity of the enzyme towards the substrate. The V_{max} of the enzyme, however, is unaffected.

Based on Fig. 1.20, the following rate equation, dissociation constants and mass balance can be written to describe competitive binding.

$$v = k_{cat}[ES] \tag{1.125}$$

$$K_s = \frac{[E][S]}{[ES]} \text{ and } K_i = \frac{[E][I]}{[EI]} \tag{1.126}$$

$$[E_T] = [E] + [ES] + [EI] = [E] + \frac{[E][S]}{K_s} + \frac{[E][I]}{K_i} \tag{1.127}$$

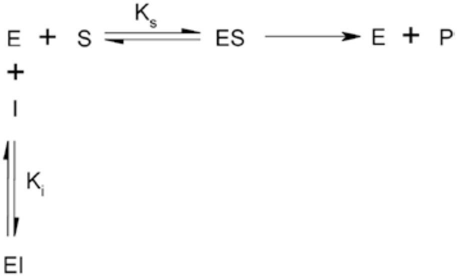


Fig. 1.20 Postulated mechanism for competitive inhibition

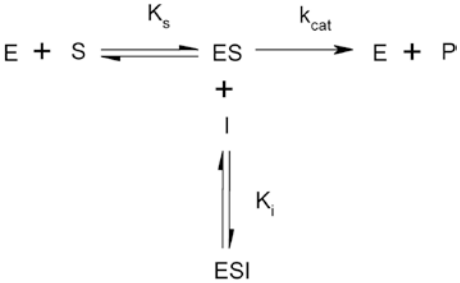


Fig. 1.21 Postulated mechanism for competitive inhibition

By dividing Eq. 1.125 by the total enzyme concentration ($\frac{v}{[E_T]}$) and rearranging, the following expression for the velocity of an enzymatic reaction in the presence of a competitive inhibitor is obtained:

$$v = \frac{V_{\max}[S]}{K_s^* + [S]} = \frac{V_{\max}[S]}{\alpha K_s + [S]} \tag{1.128}$$

where K_s^* is the apparent enzyme-substrate dissociation constant in the presence of an inhibitor. In competitive inhibition, $\alpha = 1 + \frac{[I]}{K_i}$.

1.9.8.2 Uncompetitive Inhibition

In uncompetitive inhibition, the inhibitor interacts with *only* the enzyme-substrate complex at a site other than the active site. This is illustrated in Fig. 1.21. The inhibitor does not bind to the same site on the enzyme as the substrate. Presumably, binding of the inhibitor to this site causes a conformational change that decreases the activity of the enzyme.

Uncompetitive inhibition results in an apparent decrease in both the V_{max} and K_s. The apparent increase in affinity of enzyme for substrate, i.e. a decrease in K_s, is due to unproductive substrate binding. This, in essence, results in a decrease in the free enzyme concentration. Half-maximum velocity, or half-maximal saturation, will therefore be attained at a relatively lower substrate concentration. The rate equation for the formation of product, the dissociation constants for enzyme-substrate (ES) and ES-inhibitor (ESI) complexes, and the enzyme mass balance are, respectively:

$$v = k_{cat}\left[ES\right] \tag{1.129}$$

$$K_s = \frac{[E][S]}{[ES]} \; K_i = \frac{[ES][I]}{[ESI]} \tag{1.130}$$

$$\left[E_T\right] = [E] + [ES] + [ESI] = [E] + \frac{[E][S]}{K_s} + \frac{[E][S][I]}{K_s K_i} \tag{1.131}$$

Normalization of the rate equation by total enzyme concentration ($\frac{v}{\left[E_T\right]}$) and rearrangement results in following expression for the velocity of an enzymatic reaction in the presence of an uncompetitive inhibitor:

$$v = \frac{V^*[S]}{K_s^* + [S]} = \frac{\frac{V_{\max}}{\alpha}[S]}{\frac{K_s}{\alpha} + [S]} \tag{1.132}$$

where $V_{\max}^*$ and K_s^* correspond, respectively, to the apparent enzyme maximum velocity and apparent enzyme-substrate dissociation constant in the presence of an inhibitor. In the case of uncompetitive inhibition, $V_{\max}^* = \frac{V_{\max}}{\alpha}$ and $K_s^* = \frac{K_s}{\alpha}$ where:

$$\alpha = 1 + \frac{[I]}{K_i} \tag{1.133}$$

1.9.8.3 Non-competitive Inhibition

In non-competitive inhibition, the inhibitor interacts with both the free enzyme and with the enzyme substrate complex at a site other than the active site. This is illustrated in Fig. 1.22.

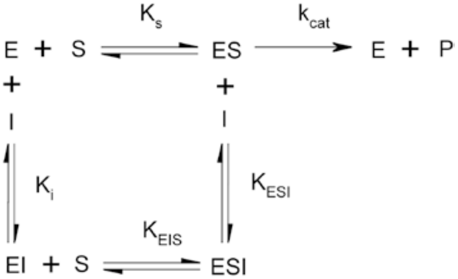


Fig. 1.22 Postulated mechanism for competitive inhibition

Unlike uncompetitive inhibition, $V_{\max}$ decreases while K_s is unaffected. The rate equations for the formation of product, the dissociation constants for enzyme-substrate (ES and ESI) and enzyme-inhibitor (EI and ESI) complexes, and the enzyme mass balance are given below:

$$v = k_{cat}\left[ES\right] \tag{1.134}$$

$$K_s = \frac{[E][S]}{[ES]} \text{ and } K_{EIS} = \frac{[EI][S]}{[ESI]} \text{ and } K_i = \frac{[E][I]}{[EI]} \text{ and } K_{ESI} = \frac{[ES][I]}{[ESI]} \tag{1.135}$$

It is assumed that the binding of either substrate or inhibitor to the free enzyme to form their respective complexes will not affect the subsequent binding of the inhibitor or substrate to these complexes. In the diagram given in Fig. 1.22, it is assumed that the binding of the enzyme-inhibitor complex (EI) to the substrate (S) has the same dissociation constant (K_s) as the binding of the substrate to the free enzyme (E). Similarly, the binding of the inhibitor to the enzyme-substrate complex (ES) is described by the same dissociation constant as the binding of the inhibitor (I) to the free enzyme (E). This mathematically translates to $K_{EIS} = K_s$ and $K_{ESI} = K_i$ to give:

$$K_s = \frac{[E][S]}{[ES]} \text{ and } K_s = \frac{[EI][S]}{[ESI]} \text{ and } K_i = \frac{[E][I]}{[EI]} \text{ and } K_i = \frac{[ES][I]}{[ESI]} \tag{1.136}$$

Normalization of the rate equation by total enzyme concentration ($\frac{v}{\left[E_T\right]}$) and rearrangement results in the following expression for the velocity of an enzymatic reaction in the presence of a non-competitive inhibitor:

$$v = \frac{V_{\max}^{*}\left[S\right]}{K_s + \left[S\right]} = \frac{\frac{V_{\max}}{\alpha}\left[S\right]}{K_s + \left[S\right]} \tag{1.137}$$

where $V_{\max}^{*}$ corresponds to the apparent enzyme maximum velocity in the presence of an inhibitor. In the case of non-competitive inhibition, $V_{\max}^{*} = \frac{V_{\max}}{\alpha}$, where:

$$\alpha = 1 + \frac{\left[I\right]}{K_i} \tag{1.138}$$

A summary of the effects of reversible inhibitors on the catalytic parameters K_s and V_{max} is presented in Table 1. The effects of these reversible inhibitors on the shape of the Michaelis-Menten curve is given in Fig. 1.23.

Table 1. Summary of the effects of reversible inhibitors on apparent enzyme catalytic parameters $V^*_{\max}$ and K^*_s

	Competitive	Uncompetitive	Noncompetitive
$V^*_{\max}$	No effect	Decrease	Decrease
K^*_s	Increase (αK_s)	Decrease ($\frac{K_s}{\alpha}$)	No effect (K_s)

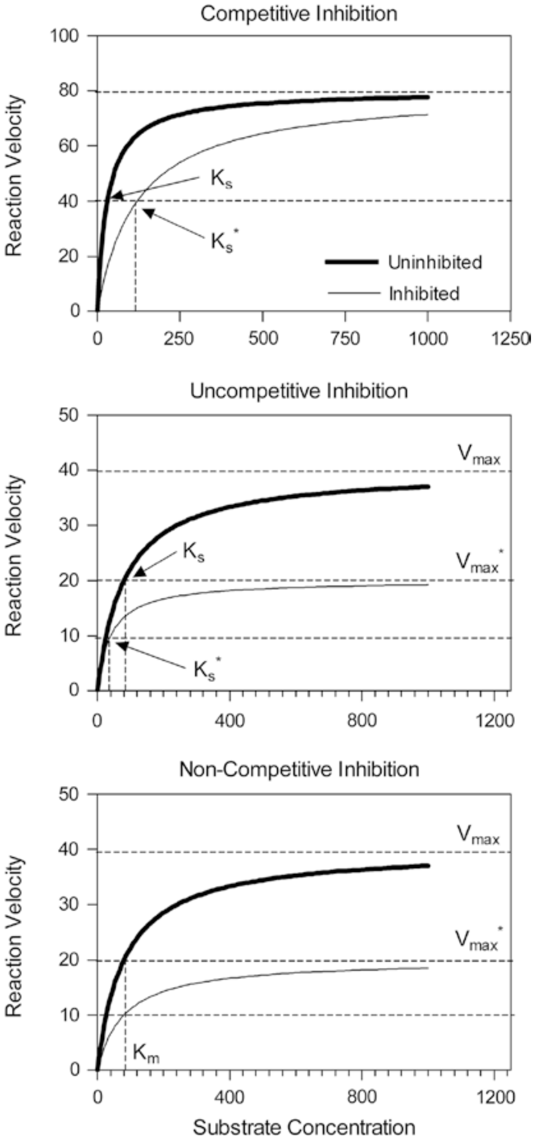


Fig. 1.23 The effects of different types of reversible inhibition on the shape of the Michaelis-Menten plot

1.9.8.4 Determining Inhibitor Affinity

A typical enzyme inhibition experiment is designed to determine the nature of the inhibition process, as well as estimate the magnitude of K_i, itself a measure of the enzyme's affinity to the inhibitor. An inhibitor with a stronger binding affinity than a given substrate will have a K_i lower than the K_s of the particular substrate. This means that a smaller concentration of inhibitor is adequate to achieve a similar degree of binding as the substrate.

To measure the K_i, the initial velocities used to generate a Michaelis-Menten plot are measured at different inhibitor and substrate concentrations. The apparent maximum reaction velocity and the apparent equilibrium dissociation constant for the enzyme-substrate complex can then be used to determine the K_i by the transformations described in Figs. 1.24a and 1.24b.

For competitive inhibition, $K_s^* = \alpha K_s$, where $\alpha = 1 + \frac{[I]}{K_i}$. Thus, $K_s^* = \left(1 + \frac{[I]}{K_i}\right) K_s$. This can be re-arranged to form:

$$K_s^* = \left(K_s + [I]\frac{K_s}{K_i}\right) \tag{1.139}$$

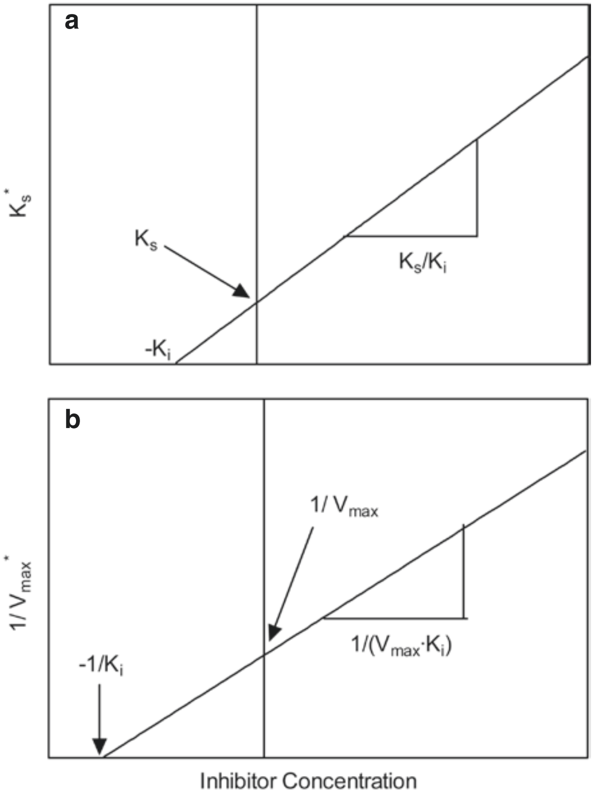


Fig. 1.24 Characteristic plots used in the determinations of the K_i for competitive inhibition (**a**), uncompetitive and non-competitive inhibition (**b**) from plots of K_s^* (**a**) and $1/V_{max}^*$ (**b**) versus inhibitor concentration, respectively

A plot of the apparent K_s (K_s^*) versus inhibitor concentration will yield a straight line with a y-intercept corresponding to K_s and a slope corresponding to $\frac{K_s}{K_i}$. This is shown in Fig. 1.24a.

For uncompetitive inhibition and non-competitive inhibition, $V_{\max}{}^* = \frac{V_{\max}}{\alpha}$ where $\alpha = 1 + \frac{[I]}{K_i}$. Thus, $V_{\max}{}^* = V_{\max} * \left(1 + \frac{[I]}{K_i}\right)^{-1}$. Taking the inverse of both sides yields

$$\frac{1}{V_{\max}{}^*} = \frac{1}{V_{\max}} * \left(1 + \frac{[I]}{K_i}\right) \tag{1.140}$$

which can be linearized to

$$\frac{1}{V_{\max}{}^*} = \frac{1}{V_{\max}} + [I]\frac{1}{V_{\max} \cdot K_i} \tag{1.141}$$

The y-intercept is the inverse of the maximum attainable reaction rate while the slope is the inverse of the product of the maximum attainable reaction rate and the K_i. This linearization is shown in Fig. 1.24b.

1.10 Food Science-Specific Kinetic Analysis: D and Z Values

A special application of the first order integrated rate equation is in the determination of 'decimal reduction times' or D values. The decimal reduction time of a reaction is the *time* required for a one $\log_{10}$ reduction in the concentration of reacting species, i.e., a 90% reduction in the concentration of reactant. Decimal reduction times are determined from the slope of $\log_{10}([A_t]/[A_o])$ versus time plots (Fig. 1.25). The modified integrated first order integrated rate equation can be expressed as:

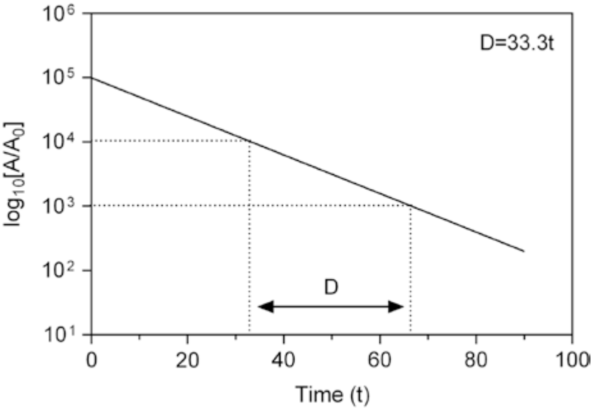


Fig. 1.25 Semi-logarithmic plot of a first order reaction showing decreases in relative concentration as a function of time. The characteristic time required for a 90% reduction in the concentration of reactant is called the Decimal Reduction Time (*D*)

$$\log_{10}\frac{[A_t]}{[A_o]} = -\frac{t}{D} \tag{1.142}$$

or

$$[A_t] = [A_o]\cdot 10^{-\frac{t}{D}} \tag{1.143}$$

The decimal reduction time (D) is related to the first order rate constant (k_r) in a straightforward fashion:

$$D = \frac{2.303}{k_r} \tag{1.144}$$

A parameter closely related to the energy of activation is the so-called “Z” value, or the temperature dependence of the decimal reduction time (D). The Z value is the temperature increase required for a one $\log_{10}$ reduction (90% decrease) in the D value, expressed as:

$$\log_{10} D = \log_{10} C - \frac{T}{Z} \tag{1.145}$$

or

$$D = C\cdot 10^{-\frac{T}{Z}} \tag{1.146}$$

where C is a constant related to the frequency factor A in the Arrhenius equation. The Z value can be determined from a plot of $\log_{10}$D versus temperature (Fig. 1.26).

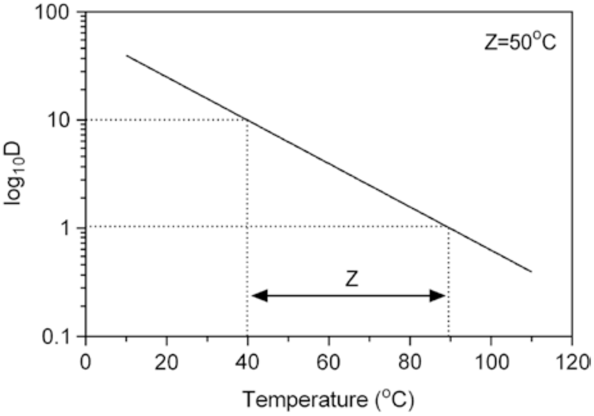


Fig. 1.26 Temperature dependence of the Decimal Reduction Time (*D*). The increase in temperature leading to a 90% reduction in D value is called the Z value

Alternatively, if D values are only known at two temperatures, the Z value can be determined using the following equation:

$$\log_{10}\frac{D_2}{D_1} = -\frac{(T_2 - T_1)}{Z} \tag{1.147}$$

It can be easily shown that the Z values is inversely related to the energy of activation:

$$Z = \frac{2.303RT_1T_2}{E_a} \tag{1.148}$$

where T_1 and T_2 are the two temperatures used in the determination of E_a.

Bibliography

1. Avery HE (1974) Basic reaction kinetics and mechanisms. The Macmillan Press Ltd, London/England
2. Dixon M, Webb EC (1979) Enzymes, 3rd edn. Academic Press, New York
3. Marangoni AG (2003) Enzyme kinetics: a modern approach. Wiley-Interscience, New York
4. Van Boekel MAJS (2009) Kinetic modeling of reactions in foods. CRC Press, Boca Raton

Part II

Kinetic Modeling of Complex Processes in Food Systems

Chapter 2
Chlorophyll Degradation in Green Tissues: Olives, Cabbage and Pickles

2.1 Chlorophyll Pigments

The chlorophylls are the main pigments in green plants, algae and other photosynthetic microorganisms. They belong to a class of pyrrole ring compounds known as porphins. Derivatives of porphins are called porphyrins. Phorbin, a porphyrin formed by the addition of another ring structure, serves as the basis for all chlorophyll molecules (Fig. 2.1). Different types of chlorophyll may arise depending on the chemical substituents on the R and R′ group (Fig. 2.2).

Chlorophylls are predominantly found in the chloroplasts of plant cells. They are of hydrophobic nature and are found alongside carotenoids, lipids and lipoproteins. The carotenoids serve an important physiological function. Carotenoids prevent

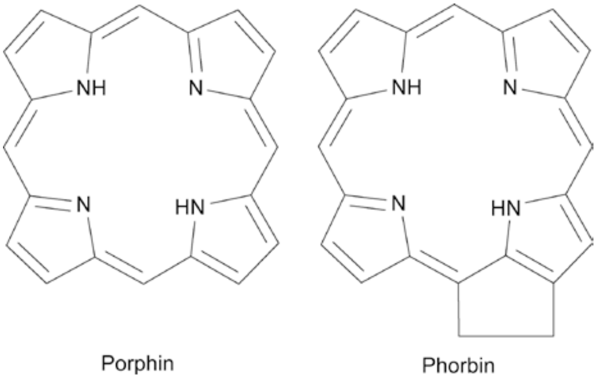


Fig. 2.1 Porphin is a compound made of 4 pyrole rings linked via methylene bridges. Phorbin is a porphin derivative with an attached 5-carbon ring structure.

A.G. Marangoni, *Kinetic Analysis of Food Systems*,
DOI 10.1007/978-3-319-51292-1_2

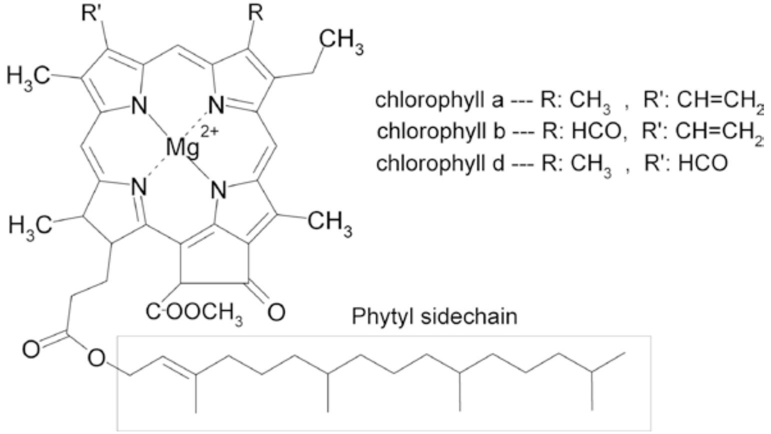


Fig. 2.2 The base molecule for chlorophyll compounds. Possible derivatives that can be obtained via the substitution of the R- and R'- groups are shown.

chlorophyll from acting as a photosensitizer in light-induced oxidation. The extraction of chlorophyll is often achieved through the use of generally non-polar solvents with some polar groups an example of which is acetone.

2.2 Chlorophyll Degradation

Chlorophyll absorbs visible light in in the wavelength range 400–500nm (blue) and 600–700nm (red). The remaining unabsorbed radiation is reflected, giving the chlorophylls their green color. When the chlorophyll molecule is altered or degraded, the absorption spectra may shift, leading to a change in color. This alteration may be natural or a result of food processing.

To the food scientist, chlorophyll loss represents a quality problem in the manufacture of processed plant products. The degradation of chlorophyll causes a shift in colour from brilliant green to olive-brown in processed foods. This results in an undesirable swamp-green appearance in foods. Chlorophyll loss, however, may serve as a useful measure of a plant's ripeness and freshness. In live plants, chlorophyll loss accompanies the natural process of senescence.

The degradation of chlorophyll results in the formation of five groups of intermediate compounds. The reason for their classification into groups rather than individual compounds is that the side chains of the chlorophyll molecule are often substituted with different groups, resulting in structurally similar, yet different compounds.

Several mechanisms contribute to the degradation of chlorophyll. The phytyl chain on chlorophyll can be cleaved off, or hydrolyzed, by the chlorophyllase enzyme to yield chlorophyllide. Chlorophyllase is dormant in live tissues but is activated by heat during processing. The enzyme exhibits a temperature optimum between 60 and 80 °C (blanching temperatures) but loses activity if it is further heated. The phytyl chain may also be removed via non-enzymatic means. The dephytylation of the chlorophyll molecule does not cause any change in the color of the pigment.

The second major chlorophyll degradation mechanism is the abstraction of the Mg^{2+} ion bound to the center of the chlorophyll molecule to yield pheophytin. Under acidic conditions, the Mg^{2+} is replaced by H^+ ions. The reaction may also be catalyzed by the magnesium dechelatase enzyme. The removal of the Mg^{2+} results in a shift of the absorption spectrum, leading to the appearance of an olive-brown, swamp-green color.

The degradation of chlorophyll to pheophorbide is a two-step process. These two steps can come in any order resulting in two distinct parallel pathways of chlorophyll degradation. In the first pathway, the phytyl chain from the chlorophyll molecule is cleaved to yield chlorophyllide. The Mg^{2+} ion bound to the center of the porphyrin ring is then replaced with H^+ ions to yield pheophorbide.

The second pathway is chemically identical to the first pathway except that the order in which the steps occur are reversed. The Mg^{2+} ion in chlorophyll is first removed by the action of either acid or magnesium dechelatase to yield pheophytin. The phytyl chain in pheophytin is then removed to yield pheophorbide. The entire reaction and the parallel nature of the steps are outlined in Fig. 2.3.

2.3 A Kinetic Model of Chlorophyll Degradation

The following rate equations can be written for the degradation of chlorophyll as given in Fig. 2.4 where A stands for chlorophyll, B stands for pheophytin, C stands for chlorophyllide and D stands for pheophorbide.

$$\frac{d[A]}{dt} = -k_1[A] - k_3[A] \tag{2.1}$$

$$\frac{d[B]}{dt} = k_1[A] - k_2[B] \tag{2.2}$$

$$\frac{d[C]}{dt} = k_3[A] - k_4[C] \tag{2.3}$$

$$\frac{d[D]}{dt} = k_2[B] + k_4[C] \tag{2.4}$$

Assuming no chlorophyll molecules are synthesized or introduced into the system, the four compounds are governed by the following mass balance:

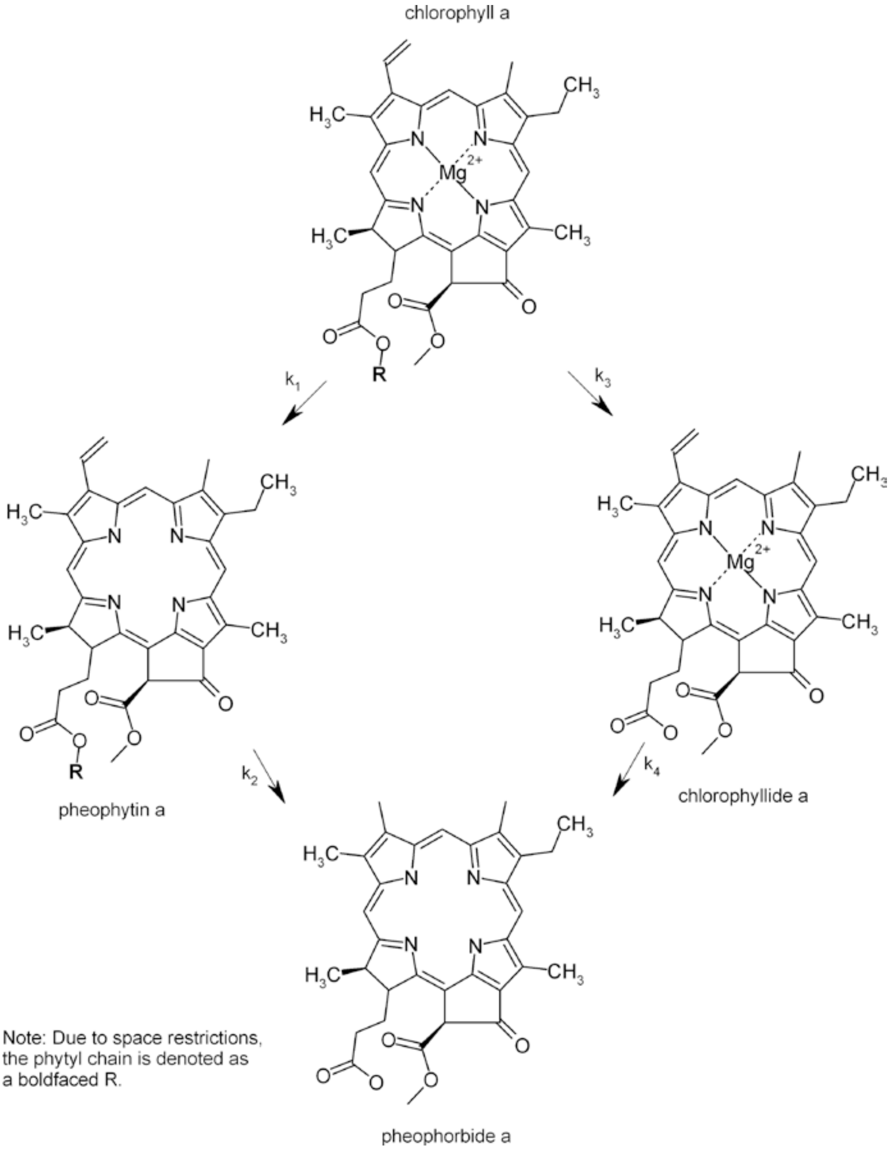


Fig. 2.3 The parallel reaction pathway for the conversion of chlorophyll *a* to pheophorbide a.

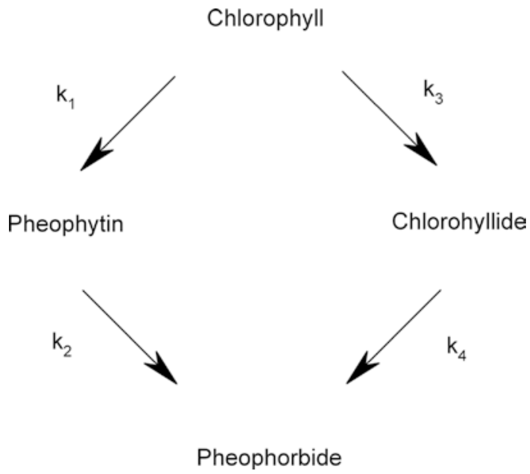


Fig. 2.4 A proposed kinetic model for the degradation of cholorophyll to pheophorbide.

$$[A_o]+[B_o]+[C_o]+[D_o]=[A]+[B]+[C]+[D] \tag{2.5}$$

where [A_o] is the initial chlorophyll concentration, [B_o] is the initial pheophytin concentration, [C_o] is the initial chlorophyllide concentration and [D_o] is the initial pheophorbide concentration. The following differential equations were solved to yield the following analytical solutions.

$$[A]=[A_o]e^{-(k_1+k_3)t} \tag{2.6}$$

$$[B]=\frac{k_1[A_o]}{k_2-k_1-k_3}\left[e^{-(k_1+k_3)t}-e^{-k_2t}\right]+[B_o]e^{-k_2t} \tag{2.7}$$

$$[C]=\frac{k_3[A_o]}{k_4-k_1-k_3}\left[e^{-(k_1+k_3)t}-e^{-k_4t}\right]+[C_o]e^{-k_4t} \tag{2.8}$$

$$[D]=[D_o]+([A_o]-[A])+([B_o]-[B])+([C_o]-[C]) \tag{2.9}$$

Instead of writing a separate differential equation describing the formation of pheophorbide, the quantity of pheophorbide was calculated by the re-arrangement of the mass balance equation. This assumes that pheophorbide is the ultimate product of chlorophyll degradation. This is a reasonable assumption since a study investigating this degradation pathway revealed that the pheophorbide concentration did not decrease even after 200 days.

The given solutions to the differential equations can be fitted to their corresponding data sets via nonlinear regression. Oftentimes, however, analytical solutions are impossible to obtain for differential equations describing more complex phenomena. A simple solution to this is to fit the data set to a numerical approximation of the differential equation. The availability of more powerful computers and sophisticated data analysis software have allowed for the fitting of data directly to differential equations instead of the analytical solutions to these same differential equations. Moreover, all equations can be fitted to all data sets simulataneously, thus creating one unique set of parameter estimates, which are more accurate and robust. This is called multiple nonlinear regression. Multiple nonlinear regression is not discussed in this book.

2.3.1 Coleslaw by Heaton et al. [2]

In a previous study establishing the theoretical framework for the kinetic modelling of chlorophyll degradation in coleslaw, Heaton and others determined that chlorophyll degradation in coleslaw followed only one particular pathway out of the two pathways available. This pathway is the *Chlorophyll* → *Pheophytin* → *Pheophorbide* pathway. The data was fitted to the model on the assumption that chlorophyll does not degrade by the other pathway available, namely *Chlorophyll* → *Chlorophyllide* → *Pheophorbide*. This was evident from the experimental data.

The experimental data were fitted using the following constraints: $k_3=0$, $k_4=0$, $A_{lim}=6$ mol%, $B_{lim}=20.5$ mol%, $C_{lim}=0.1$ mol%, $A_o=58.4$ mol%, $B_o=27.9$ mol%, $C_o=0.1$ mol%, $D_o=13.6$ mol%. The assumption that chlorophyll only follows the *Chlorophyll* → *Pheophytin* → *Pheophorbide*pathway is embodied in the model by setting k_3 and k_4 equal to 0. A_{lim}, B_{lim} and C_{lim} denote the limiting concentration of chlorophyll, pheophytin and chlorophyllide, respectively. The limiting concentration denotes the minimum amount of the substance that exists in the system. Presumably, the limiting concentration is the concentration of each particular intermediate that is unavailable for degradation but is nevertheless present in the system. The difference between the reactant's concentration (for example, $[A]$) and its limiting concentration A_{lim}, $[A]-[A_{lim}]$, gives the concentration of reactant available to undergo degradation.

The calculated curves for this model and the experimental points are presented in Fig. 2.5, demonstrating that the model described the experimental results quite accurately.

2.3.2 Pickles by White et al. [6]

To ensure that the given differential equations also applied to chlorophyll degradation in other food systems, the model was used to describe chlorophyll degradation in brined cucumbers (pickles). White et al. noted that the pheophytin levels in the system studied remained constant after an initial increase early on in the experiment. This observation translates to $k_2=0$, that is, pheophytin is not

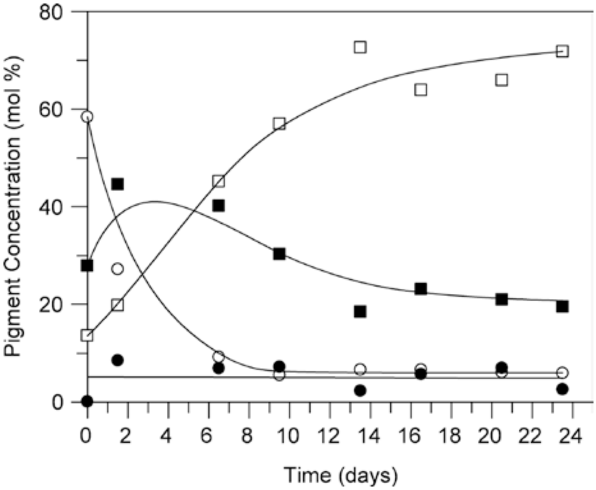


Fig. 2.5 Chlorophyll degradation data for processed coleslaw. (○) chlorophyll; (●) chlorophyllide; (■) pheophytin; (□) pheophorbide. The solid lines represent the curves obtained by regressing the analytic solutions to the data (Data by Heatonet al. [2]).

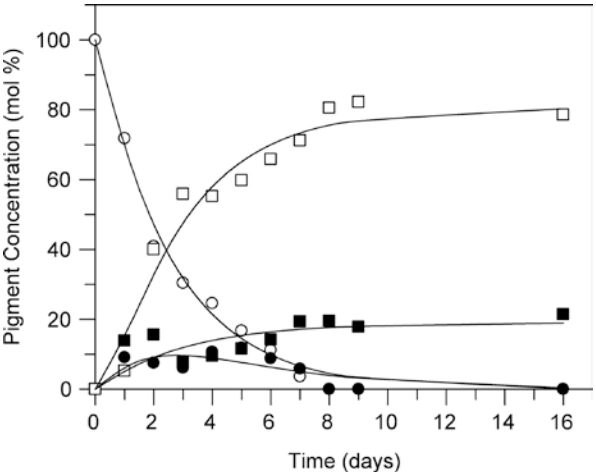


Fig. 2.6 Chlorophyll degradation data for brined cucumbers. (○) chlorophyll; (●) chlorophyllide; (■) pheophytin; (□) pheophorbide. The solid lines represent the curves obtained by regressing the analytic solutions to the data (Data by White et al. [6]).

further degraded to pheophorbide and that pheophytin accumulates in the food system, reaching a constant concentration towards the end of the observation period.

White's experimental data was fitted with the following boundary conditions: $k_2=0$, $A_o=100$ mol%, $B_o=0$, $C_o=0$, $D_o=0$, $A_{lim}=0$, $B_{lim}=21.4$ mol%, $C_{lim}=0$, $D_{lim}=78.6$ mol%. The experimental data and the resulting model curve is given in Fig. 2.6. The good fit observed implies that the model is also applicable to describe chlorophyll degradation in brined cucumbers.

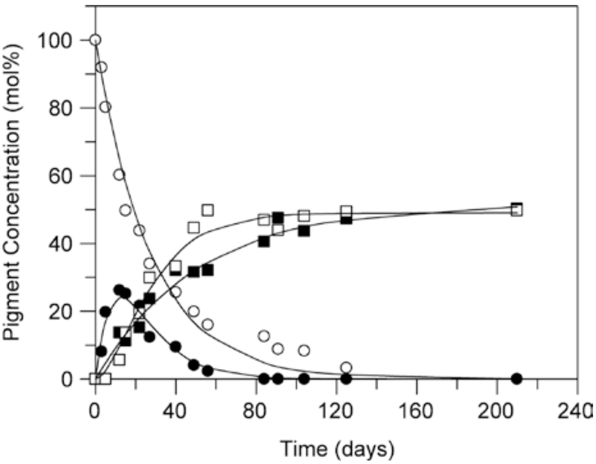


Fig. 2.7 Chlorophyll degradation data for whole brined olives. (○) chlorophyll; (●) chlorophyllide; (■) pheophytin; (□) pheophorbide. The solid lines represent the curves obtained by regressing the analytic solutions to the data (Data by Minguez-Mosquera et al. [4]).

2.3.3 Olives by Minguez-Mosquera et al. [4]

To further test the utility of the model, experimental data on the degradation of chlorophyll in processed olives was fitted to the given model. Minguez-Mosquera et al. proposed a similar pathway of chlorophyll degradation as that proposed by White. Fig. 2.7 shows the experimental data and the modelled curve. Once again, the closeness-of-fit indicates that the model is also valid for brined olives.

2.3.4 Relating Kinetic Parameters to Degradation Mechanisms

The kinetic parameters derived from fitting the analytical solutions of the model to the data is given in Table 2.1. Consider the degradation of *Chlorophyll* → *Pheophytin*. It is believed that this reaction occurs rapidly under acidic conditions. At low enough pH, H^+ ions can displace the Mg^{2+} ion chelated to the porphyrin ring. A glance at the k_1 constants of the different data sets reveal that the formation of pheophytin from chlorophyll was fastest in coleslaw ($k_1 = 0.54$ day^{-1}), followed by the pickles ($k_1 = 0.084$ day^{-1}) and then by the olives ($k_1 = 0.023$ day^{-1}). The vastly greater k_1 constant observed in the early portion of the coleslaw experiment is due to the almost instantaneous drop in the pH of coleslaw, presumably caused by the addition of salad dressing.

Similarly, no change was observed in the chlorophyllide concentration in coleslaw since the beginning of the experiment. This is the basis for the abovementioned assumption $k_3 = 0$ and $k_4 = 0$ in Heaton *et al*'s data. However, the degradation of chlorophyll to chlorophyllide was the primary initial pathway observed in brined olives ($k_3 = 0.033$ day^{-1}) and pickled cucumbers ($k_3 = 0.29$ day^{-1}). A likely explanation as to the high rate of conversion from *Chlorophyll* → *Chlorophyllide* (relative to

Table 2.1 Kinetic constants (mean ± standard deviation) derived from fitting the individual analytical solutions of the model to the data

Food system	k_1 (day^{-1})	k_2 (day^{-1})	k_3 (day^{-1})	k_4 (day^{-1})
Pickled cucumbers (White et. al)	0.084 ± 0.015	0	0.29 ± 0.01	1.12 ± 0.70
Brined olives (Minguez-Mosquera et. al)	0.023 ± 0.005	0	0.33 ± 0.013	0.056 ± 0.020
Coleslaw (Heaton et. al)	0.54 ± 0.06	0.22 ± 0.05	0	0

the corresponding rates of conversion to pheophytin) is the high activity of the chlorophyllase enzyme in the unprocessed olives. Prior to brining and fermentation, the pHs of both the cucumber and the olives are neutral. Chlorophyllase has high activity at neutral pH. Furthermore, the pH of the olives and the cucumbers remained at the neutral optimum for several days --- ample time to degrade chlorophyll to chlorophyllide. Subsequent fermentation results in a decrease in pH, completing the degradation pathway by converting chlorophyllide to pheophorbide. This decrease in pH also resulted in the conversion of chlorophyll to pheophytin as observed in Figs. 2.6 and 2.7.

Bibliography

1. Heaton JW, Marangoni AG (1996) Chlorophyll degradation in processed foods and senescent plant tissues. Trends Food Sci Technol 7:8–15
2. Heaton JW, Yada RY, Marangoni AG (1996) Discoloration of coleslaw is caused by chlorophyll degradation. J Agric Food Chem 44:395–398
3. Heaton JW, Lencki RW, Marangoni AG (1996) Kinetic model for cholorphyll degradation in green tissue. J Agric Food Chem 44:399–402
4. Minguez-Mosquera MI, Garrido-Fernandez J, Gandul-Rojas B (1989) Pigment changes in olives during fermentation and brine storage. J Agric Food Chem 37:8–11
5. Van Boekel MAJS (2000) Anniversary Review: kinetic modeling in food science: a case study on chlorophyll degradation in olives. J Sci Food Agric 80:3–9
6. White RC, Jones ID, Gibbs E (1963) Determination of chlorophylls, chlorophyllides, pheophytins and pheophorbides in plant material. J Food Sci 28:431–435

Chapter 3
Oil Migration Through Cocoa Butter: The Solubilization-Recrystallization-Diffusion (SRD) Model

3.1 Oil Migration in Confectionery Products

Oil migration is a serious quality problem for the food industry and is a leading factor in the deterioration of most fat - structured products. Chocolate enrobed products (e.g. chocolate-coated biscuits, filled chocolate bars or shells, wafers with fat-based cream fillings and, pralines) are particularly vulnerable to this phenomenon and it causes significant softening of chocolate coatings, hardening of fillings, bloom formation and a corresponding decrease in consumer acceptability. Many studies have already tested the effects of storage conditions and variations in filling and chocolate compositions; however, the exact mechanism of oil migration still remains hypothetical and controversial. A diffusion model, commonly used in biology and medicine to define mass transfer, has also been applied to food systems in order to explain mass transport processes. However, considering the heterogeneous characteristics of certain food media such as chocolate (mixture of cocoa, sugar and milk particles in a fat matrix), it is conceivable that other mechanisms like capillary flow, matrix dissolution and recrystallization may also play a role in oil migration.

3.2 Proposed Oil Migration Mechanisms

Oil migration in confectionery type matrices has been studied for more than 50 years and its mechanism is still subject to controversy.

A.G. Marangoni, *Kinetic Analysis of Food Systems*,
DOI 10.1007/978-3-319-51292-1_3

3.3 Fickean Diffusion

Ziegleder and coworkers first explained oil migration mechanisms in 1996 using Fickean diffusion. They started with a solution of Fick's second law of the form:

$$\frac{c_t}{c_0} = \frac{2K \cdot A\sqrt{D \cdot t}}{V \cdot \sqrt{\pi}} \tag{3.1}$$

with c_t and c_0 being the concentration of migrating species at time t and at $t = 0$, A is the contact area between the phases, V is the volume of the fat phase through which migration is taking place, D is the diffusion coefficient and t is the migration time. Here, K describes the contact between the two phases and $K = 1$ for perfect, $K<1$ for insufficient contact and $K>1$ for structural changes due to swelling or eutectic effects.

Ziegleder then proposed a further simplified form of the equation based on a number of assumptions:

$$\frac{m_t}{m_\infty} = \frac{A\sqrt{D \cdot t}}{V} \tag{3.2}$$

where m_t represents the mass migrated at time t and m_∞ the final mass of migrating material at equilibrium.

They suggested that in the short-time limit, m_t/m_∞ increases linearly with $\sqrt{t}$ and thus the apparent diffusion coefficient can be calculated from the slope of the curve as shown in Fig. 3.1. We should stop at this point and discuss the differences

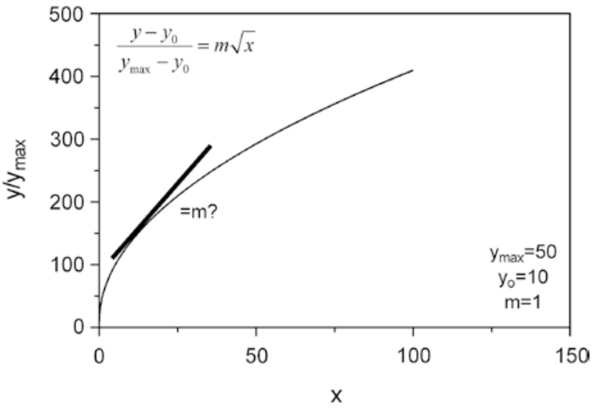


Fig. 3.1 Simulation of the generalized form of the diffusion oil migration model commonly used in the literature. The diffusion coefficient is usually determined from the slope in the "linear" initial region of the plot

between the figure reported in his paper of 1996 and the one generated by our simulation (Fig. 3.1) since they are not the same. As a matter of fact, the graph of the function shown as Eq. 3.2 bears little resemblance to the one reported originally in 1996. It is very difficult to identify a linear region in this curve. This is somewhat problematic since the slope of this "linear" region, according to Ziegleder, is required to determine the diffusion coefficient for the migration process. Moreover, there are other serious problems with this equation. For example, it does not plateau at m_∞. Our value of y_{max} was set at 50 for the simulation, but the curve just keeps increasing as a function of increasing x, beyond $y = 50$. It is not clear either how one goes from Eqs. 3.1 and 3.2. The use of this popular model should, probably, be abandoned.

Even though this model has been extensively used to describe oil migration in cocoa butter and chocolate by many researchers over the years, including our group, it does not take into account the complex microstructure of chocolate and the very specific characteristics of fat (crystal network structure, polymorphism, melting and crystallization behavior, fat matrix dissolution, limited amount of migrating oil, recrystallization). Diffusion processes are probably likely to occur, but the type of diffusion might differ from a Fickean one. This behavior us expressed as a t^n type time dependence, where $n \neq 0.5$. Thus, non-negligible and time-dependent physical modifications of the dark chocolate matrix have to be considered in order to explain the transport mechanisms and obtain accurate estimates of diffusion coefficients. In addition, the similarity in the analytical solutions of both Fickean diffusion and capillary flow models (both express a square root of time dependence) needs to be at least considered.

3.4 Capillary Forces

Chocolate can be described as a dense medium containing solid particles of cocoa, milk and sugar dispersed in a continuous crystallized fat phase. Within this matrix, pores may exist and their role in oil migration should also be considered. From a mathematical point of view, the Lucas-Washburn equation is used to model flow and assuming the capillary is cylindrical and in contact with an infinite liquid reservoir:

$$\frac{2}{r}\gamma\cos\theta = \frac{8}{r^2}\mu h\frac{dh}{dt} + \rho g h \tag{3.3}$$

where h is the distance the fluid is drawn into the capillary, γ is the surface tension of the fluid, θ is the contact angle between the fluid and the capillary wall, ρ and μ are the density and viscosity of the liquid, respectively, g the acceleration due to the gravity and r the radius of the pore.

The solution to an ordinary differential equation of the form $\frac{dh}{dt}+\frac{a}{h}-b=0$ in the asymptotic limit of long times can be approximated by an exponential function of the form:

$$h = h_{\infty}\left[1-\exp\left(-\alpha t\right)\right] \tag{3.4}$$

where h_{∞} is the height at equilibrium reached by the liquid within the capillary. The coefficient α has the following form:

$$\alpha = \frac{\rho g r^2}{8\mu h_{\infty}} \tag{3.5}$$

At short times, and in a horizontal capillary, the analytical solution of Eq. 3.3 predicts a linear relationship between h and $\sqrt{t}$ as described by:

$$h = \sqrt{\frac{r\gamma\cos\theta}{2\mu}t} \tag{3.6}$$

This approximation was validated by curve fitting in Fig. 3.2. A simulated exponential association curve (Eq. 3.4) was fitted to a function with a square root dependence of time and the fit had an r^2 of 0.99, which was acceptable.

The principal limitations in using this model are mainly a lack of internal microstructural data for chocolate and the difficulty in determining the contact angle between the capillary wall and the liquid, the viscosity and the surface tension of the migrating fluid. For more information on capillary issues, the reader should read the work from Rousseau's group listed in the bibliography.

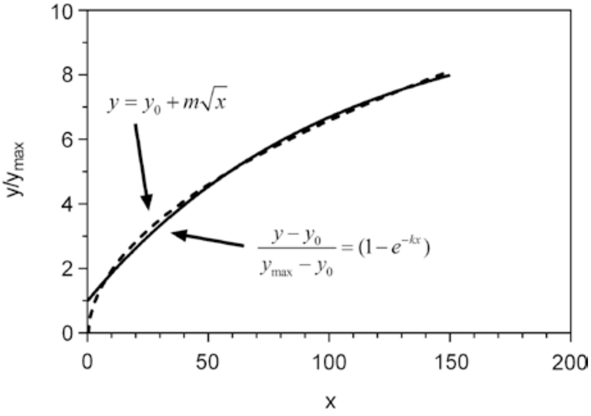


Fig. 3.2 Simulation of the generalized form of the capillary flow migration model showing how a square root of time function approximates a one phase exponential association model in the early stages of flow

3.5 Other Considerations

From previous studies, it is clear that diffusion and capillary flow might explain most of the oil migration behavior within confectionery matrices. However, "secondary" type phenomena induced by TAG migration such as matrix dissolution or crystal modification via Ostwald ripening should also be considered.

It is also important to consider the system being studied. The mechanisms responsible for oil migration in single-phase or multiphase systems will differ. In plain chocolates, assuming homogeneous fat, cocoa powder, sugar and milk powder distributions, TAG concentration gradients do not exist and thus the main mechanism for oil transport will be capillary movement through capillaries in the chocolate. Thus, the micro- and nano-structural porosity of chocolate network will play a major role. In the presence of migrating oil, dissolution of small crystals to the benefit of larger ones via Ostwald ripening and solid-liquid polymorphic transitions (due to the lower free energy of Form VI) will also take place. Diffusional transport of TAGs from dissolved crystals will also be observed. Furthermore, since the coefficient of thermal expansion is different between solid and liquid phases, dissolution and re-crystallization will simultaneously induce the formation of further cracks or pores due to repeated contraction and expansion, and thus possibly further enhance capillary transport. On the other hand, this contraction-expansion process combined with transient overpressure forces created by a higher volume fraction of liquid oil (occupying a greater volume than solid lipid) could even lead to the convective flow of oil; a microstructural pump. This will be accentuated during temperature cycling or high temperature storage.

Conversely, in filled type chocolate, diffusion due to TAG concentration gradients will play a major role in oil migration, though capillary forces may likely also take place due to the same reasons as explained above. In this case, fat phase interactions will be critical and phase diagram determination will be necessary to better understand/predict the overall structure modification. For multi-phase systems, the most common causes of enhanced oil migration include (a) the presence of incompatible fats (e.g. lauric type fats vs. SOS ones) leading to the formation of eutectic mixtures with cocoa butter, (b) the migration of oil, like hazelnut oil, into chocolate leading to the dissolution of solid fat, as it will act as a solvent.

Triglyceride phase behavior is generally well understood, however, the role of TAG dissolution on oil migration is not. Dissolution of solid triglycerides into migrating oil moving due to diffusion or capillary action will induce matrix softening, which in turn will enhance further movement by decreasing medium viscosity or creating new pores through which oil can move. These effects will have a strong temperature dependence.

In what follows, we propose a new model for describing oil migration in confectionery products, namely a simple model that takes into account solubilisation, recrystallization and diffusion/capillary movement.

3.6 Solubilization-Recrystallization-Diffusion Model

One common thread in all oil-migration work is that the models used are too simple to describe the complexity of such a process. This complexity can be appreciated from looking at the swelling of a cylindrical piece of cocoa butter in contact with an oily "cream" for several months at room temperature (Fig. 3.3). The amount of swelling was dramatic. This casts doubt unto the validity of, for example, the value of diffusion coefficients determined by using the simple models described above. As stated above, many authors have mentioned the possibility of dissolution playing an important role in the oil migration saga. We decided to verify how much dissolution actually takes place when you drop some solid cocoa butter into peanut oil (2:8 w/w) at 18 °C and incubated statically for a long time.

Figure 3.4 shows the dynamics of dissolution of tempered and non-tempered cocoa butters from different geographical origins in peanut oil (EC-Ecuador, Brazil-BR, Malaysia-MA, China-CH, Ivory Coast- IV, Nigeria-Ni). The dissolution was, surprisingly, quite significant. In order to quantify the amount and rate at which dissolution was taking place, we fitted the data to a single exponential decay model, obtaining estimates for the rate constant of dissolution (k_S) and the span of dissolution (M), i.e., the total amount of solids dissolved, or drop in solid fat content, SFC (Table 3.1).

One conclusion form this work is that cocoa butter does dissolve in liquid vegetable oil, and that most non-tempered cocoa butter samples dissolved more than tempered ones. What was very surprising, though, is that tempered samples dis-

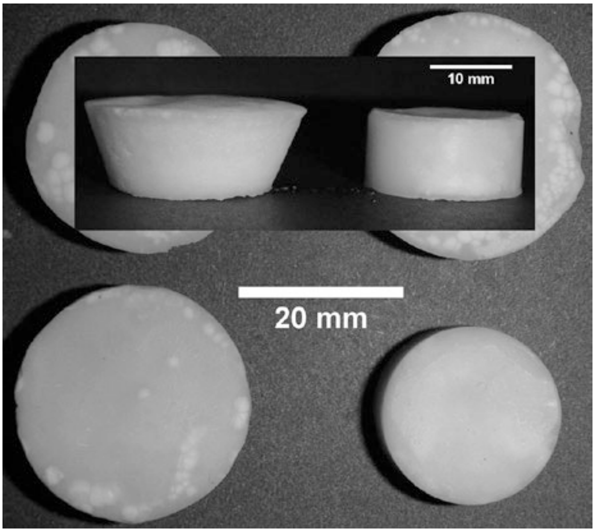


Fig. 3.3 Changes in the volume and shape of a cocoa butter cylinder exposed to an oil-rich cream phase for a period of a year at room temperature. The cylinder was placed on top of the cream layer. The side in contact with the oil phase is the smaller one

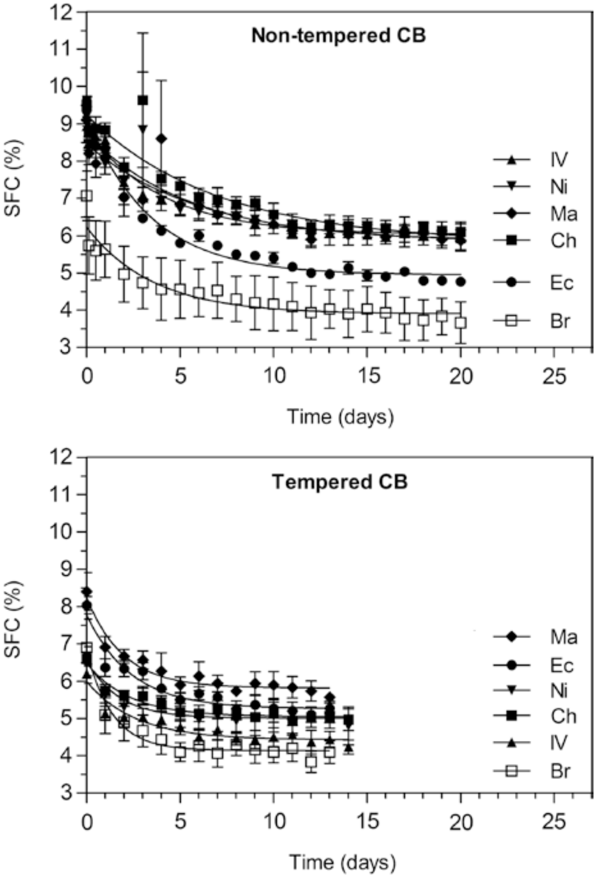


Fig. 3.4 Solubilization-induced decreases in the solid fat content as a function of time for a 20:80 (w/w) mixture of cocoa butter in peanut oil at 18 °C. The different cocoa butters are from different geographical origins *Iv* Ivory Coast, *Ni* Nigeria, *Ma* Malaysia, *Ec* Ecuador, *Ch* China, *Br* Brazil

solved faster than non-tempered ones, but, as stated above, less cocoa butter dissolved.

Thus, these results suggested that dissolution was indeed possible in an oil migration process. One could envision the migrating liquid oil interacting with the cocoa butter matrix, dissolving it and moving further. This could be taking place possibly after an initial rapid movement of the oil by capillary action through small pores, and a slower diffusive movement through the matrix. However, if dissolution is possible, then recrystallization of the dissolved and migrating fat molecules is also possible.

Thus, we decided to create a simple mathematical model that would encompass this dynamic. The model is shown diagrammatically in Fig. 3.5. The logic behind this simple model will be described next.

Table 3.1 First order dissolution rate constants and total amount of cocoa butter solids mixed into peanut oil (20:80 w/w) at 18 °C, as determined by solid fat content measurements as a function of time by pulse NMR

	M (%SFC)	k_S (day^{-1})
ECT	2.5 (0.19)	0.42 (0.074)
ECNT	4.2 (0.094)	0.29 (0.018)
BRT	2.6 (0.37)	0.72 (0.22)
BRNT	2.3 (0.37)	0.28 (0.12)
MAT	2.4 (0.27)	0.51 (0.13)
MANT	2.6 (0.30)	0.16 (0.055)
CHT	1.4 (0.16)	0.44 (0.11)
CHNT	3.4 (0.34)	0.14 (0.041)
IVT	1.5 (0.20)	0.38 (0.11)
IVNT	2.8 (0.090)	0.24 (0.022)
NIT	1.5 (0.29)	0.57 (0.23)
NiNT	2.9 (0.27)	0.18 (0.050)

Values represent the least square fit estimates and standard errors of $n = 42–45$ data points, depending on the data set *EC* ecuador, *BR* Brazil, *MA* Malaysia, *CH* China, *IV* Ivory Coast, *Ni* Nigeria, *T* tempered, *NT* non-tempered

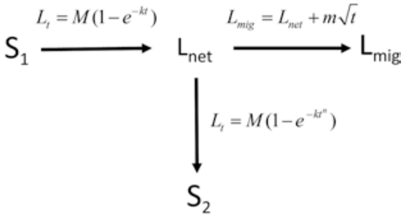


Fig. 3.5 Solubilization-Recrystallization-Diffusion model for modelling and predicting oil migration through cocoa butter

Cocoa butter comes into contact with migrating oil and dissolves following a first order process characterized by a rate constant, k_S, and with a total “mobile” fraction of triglycerides (TAGs), M. The increase in “free-to-migrate” liquid molecules is thus characterized by

$$L_t = M\left(1 - e^{-k_s t}\right) \tag{3.7}$$

These liquid TAGs, or a fraction therefrom, can re-crystallize back into their original solid form (S_1), or into another solid (S_2). This process would take place in an Avrami-like fashion (Chap. 6) characterized by a rate constant of crystallization (k_c) and an Avrami index (n):

$$S_t = M(1 - e^{-k_c t^n}) \tag{3.8}$$

The net amount of liquid TAGs which are ready to migrate can be estimated from the difference between these two processes, namely $L_t - S_t = L_{net}$, and the migration of oil (L_{mig}) is assumed to be of diffusive and/or capillary in nature and thus:

$$L_{mig} = L_{net} + m\sqrt{t} \tag{3.9}$$

Processes characterized by a square root of time dependence are usually diffusive in nature, while m is a phenomenological constant representing a migration "oil migration rate constant". However, as shown in Eq. 3.6, this type of square-root-of-time dependence could be due to capillary action. Introducing L_{net} into Eq. 3.9 using the expressions in Eqs. 3.7 and 3.8, yields

$$L_{mig} = M\left(1 - e^{-k_s t}\right) - M(1 - e^{-k_c t^n}) + m\sqrt{t} \tag{3.10}$$

which conveniently simplifies to the final form of this analytical solution:

$$L_{mig} = M\left(e^{-k_c t^n} - e^{-k_s t}\right) + m\sqrt{t} \tag{3.11}$$

Please note that we are assuming that the mobile fraction, M, is the maximum amount of TAGs that can both dissolve and recrystallize in this process.

This simple little function proved to be very powerful and fit the data incredibly well (Fig. 3.6). Reversals, nonlinearities, plateaus were all modelled very well with this function and produced a series of very interesting results from the fits (Table 3.2).

One should remind oneself that a good fit to the data does not mean that the model is correct. Fits of the model to the data were excellent for non-tempered cocoa butters (*NT*), with a Pearson correlation coefficient, r^2>0.96. On the other hand, due to the data scatter in the tempered cocoa butter (*T*) samples, r^2 values ranged from 0.33 to 0.91.

Some very interesting insights became apparent from this simple modelling exercise. Firstly, the amount of mobile TAGs (M) is always much greater in the non-tempered samples than in the tempered samples. This makes logical sense. Also, the rate of oil migration by diffusion (m) is always greater for non-tempered than for tempered samples. Finally, the rate of dissolution (k_s) is greater for non-tempered samples than for tempered samples, except for cocoa butter from Ivory Coast, where they are similar. On the other hand, behaviors for the rate of recrystallization (k_c) and Avrami index (n) were not consistent and thus trends were not evident; sometimes tempered ones were higher, sometimes non-tempered ones were. So, one clear conclusion form this work is that a reasonable strategy to mitigate oil migration in chocolate could be a reduction in the mobile fraction of TAG molecules in CB. These mobile TAGs should be identified and eliminated by breeding, or genetically modifying, the cacao plant to yield fruit with a cocoa butter devoid of these troublesome "mobile" TAGs.

As a final analysis, we compared two *tempered* cocoa butter samples with the highest (Nigeria) and lowest (China) oil migration rates (Fig. 3.7). Inspection of Table 3.2 with hopes of finding a simple explanation for these extremes will prove disappointing, since a clear trend was not identified within tempered samples.

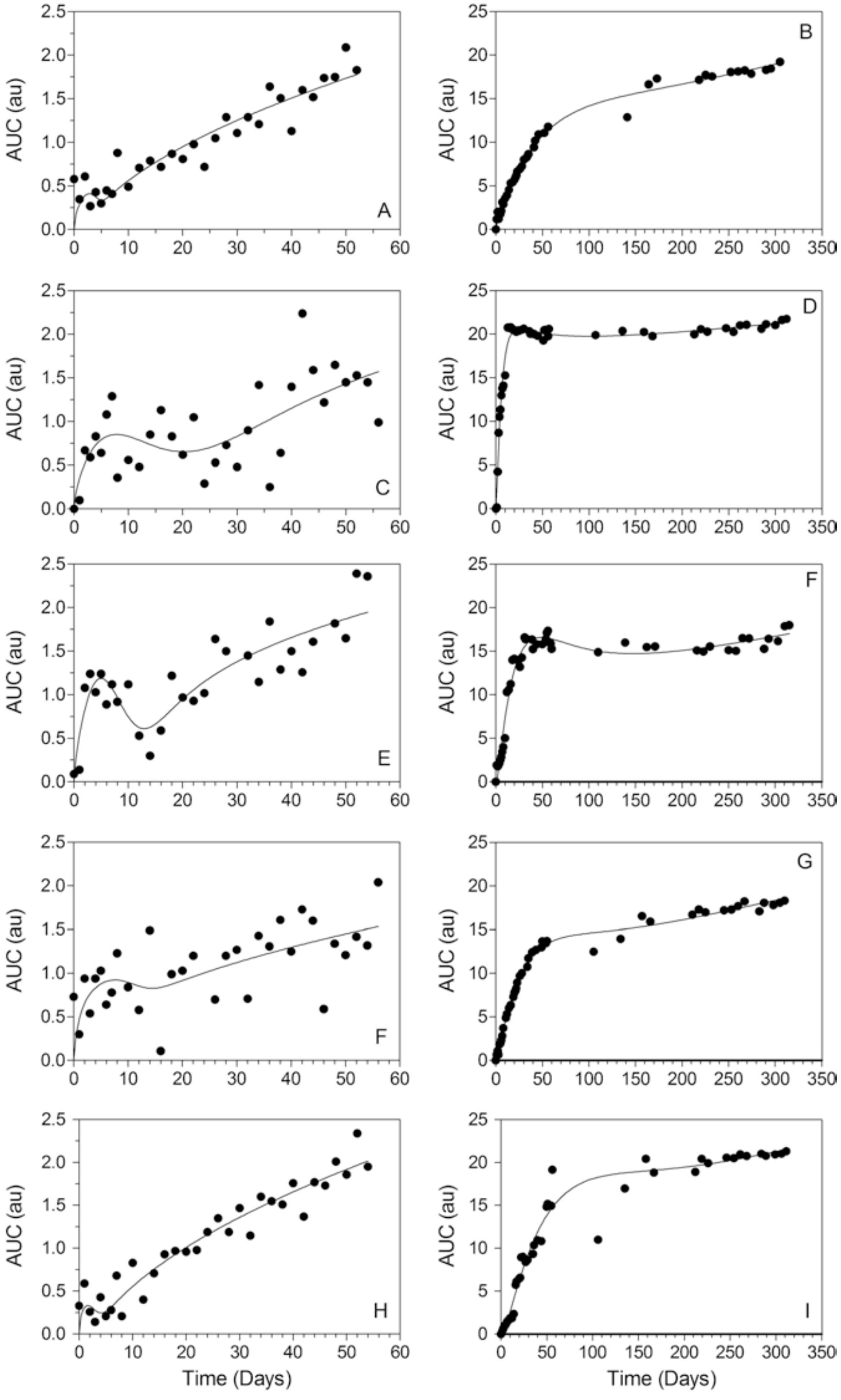


Fig. 3.6 Fits of the SRD model to oil migration data for tempered and non-tempered cocoa butters of different geographical origins

Table 3.2 Parameter estimates for the solubilisation-recrystallization-diffusion (SRD) oil migration model developed in this work

	M	k_S	n	k_C	m
ECT	0.35	0.0043	3.8	0.0048	0.28
ECNT	34	0.029	0.88	0.022	1.1
BRT	2.6	0.0076	1.8	0.055	0.23
BRNT	29	0.16	0.37	0.18	0.79
MAT	2.4	0.0033	2.6	0.12	0.26
MANT	39	0.093	0.65	0.061	0.91
CHT	0.45	0.00014	3.6	0.84	0.21
CHNT	26	0.073	0.68	0.041	1.1
IVT	0.54	0.069	2.3	0.0013	0.34
IVNT	68	0.048	0.81	0.024	1.2
NIT	1.9	0.000093	3.9	0.011	0.59

Oil migration was determined as described in [13]

As a modeller, one needs to realize the limitation of the approach. Firstly, the experimental error in the migration data is high for the tempered samples, mainly because the amount of migration is very small. Thus, it is difficult to trust values of the model parameters obtained when comparing small differences between samples with very little oil migration. Also, it is possible that migration actually involves all factors listed in combination and not only a single controlling one. On the other hand, among non-tempered samples, the highest value of *m* was associated with the highest value of *M* in the data set for cocoa butter from Ivory Coast, while the lowest value of *m* was associated with the second lowest value of *M* in the data set for Brazilian cocoa butter.

Thus, considering the limitations in the resolution of the model and data, we can ascertain that decreasing the amount of mobile TAGs in cocoa butter constitutes the best strategy to mitigate oil migration in confectionery products. Obviously, tempering has a greater effect on oil migration than cocoa butter geographical origin. Improper tempering will lead to greater oil migration and associated ensuing quality problems.

Here we have presented a simple and interesting kinetic modelling exercise on oil migration through cocoa butter from different geographical origins. This exercise has yielded valuable insights into the possible factors that influence oil migration through confectionery products.

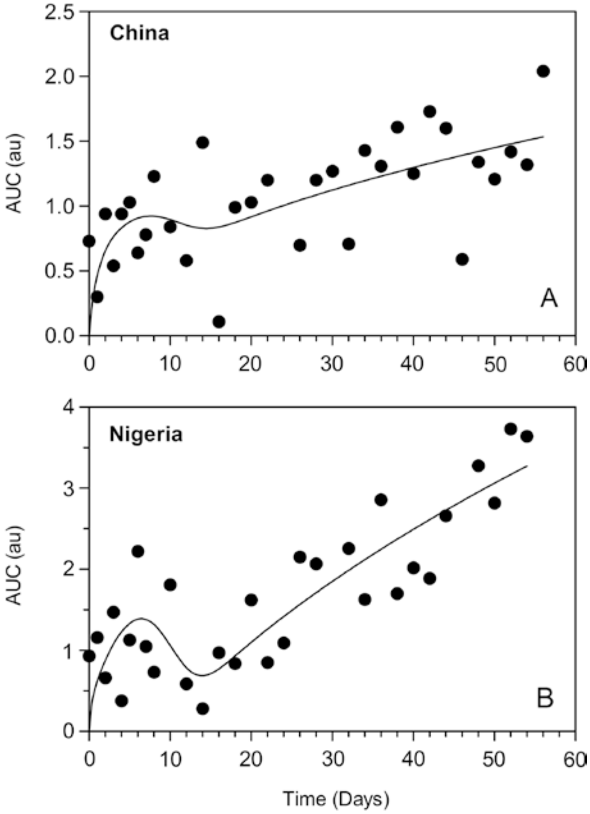


Fig. 3.7 Fits of the SRD model to oil migration data for tempered Chinese and Nigerian cocoa butter

Bibliography

1. Acevedo NC, Macmillan B, Newling B and Marangoni AG (2016) Shear effects on the diffusive movement of oil in triacylglycerol networks RSC Advances Y: xxx-xxx
2. Aguilera JM, Michel M, Mayor G (2004) Fat migration in chocolate: diffusion or capillary flow in a particulate solid? A hypothesis paper. J Food Sci 69:R167–R174
3. Altan A, Lavenson DM, McCarthy MJ, McCarthy KL (2011) Oil migration in chocolate and almond product confectionery systems. J Food Sci 76(6):E489–E494
4. Altimiras P, Pyle P, Bouchon P (2007) Structure-fat migration relationships during storage of cocoa butter models-bars: bloom development and possible mechanism. J Food Eng 80:600–610
5. Deka K, MacMillan B, Ziegler GR, Marangoni AG, Newling B, Balcom BJ (2006) Spatial mapping of solid and liquid lipid in confectionery products using a 1D centric SPRITE MRI technique. Food Res Int 39:365–371
6. Green NL, Rousseau D (2015) Oil diffusivity through fat crystal networks. Soft Matter 11:5523–5530
7. Khan RS, Rousseau D (2006) Hazelnut oil migration in dark chocolate - kinetic, thermodynamic and structural considerations. Eur J Lipid Sci Technol 108:434–443
8. Maleky F, Marangoni AG (2011) Nanoscale effects on oil migration through triacylglycerol polycrystalline colloidal networks. Soft Matter 7:6012–6024

9. Maleky F, McCarthy KL, McCarthy MJ, Marangoni AG (2012) Effect of cocoa butter structure on oil migration. J Food Sci 77:E75–E79
10. Marty S, Baker K, Neves Rodrigues J, Dibildox-Alvarado E, Marangoni AG (2005) Imaging and quantification of oil migration through cocoa butter using a flat-bed scanner. Food Res Int 38:1189–1197
11. Marty S, Baker K, Marangoni AG (2009) Optimization of a scanner imaging technique to accurately study oil migration kinetics. Food Res Int 42:368–373
12. Marty S, Schroeder M, Baker KE, Mazzanti G, Marangoni AG (2009) Small molecule diffusion through polycrystalline triglyceride networks quantified using fluorescence recovery after photobleaching. Langmuir 25:8780–8785
13. Marty S, Marangoni AG (2009) Effect of cocoa butter origin, tempering procedure and structure on oil migration kinetics. Cryst Growth Des 9:4415–4423
14. Tanuj M, Hanselmann W, Ramaswamy A (2011) Diffusion, counter-diffusion and lipid phase changes occurring during oil migration in model confectionery systems. J Food Eng 104:186–195
15. Rousseau D, Smith P (2008) Microstructure of fat bloom development in plain and filled chocolate confections. Soft Matter 4:1706–1712
16. Rousseau D (2006) On the porous mesostructure of milk chocolate viewed with atomic force microscopy. Lebensm Wiss Technol 39:852–860
17. Ziegleder G, Moser C, Geier-Greguska J (1996a) Kinetik der Fettmigration in Schokoladenprodukten Teil I: Grundlagen und Analytik. Fett-Lipid 98:196–199
18. Ziegleder G, Moser C, Geier Greguska J (1996b) Kinetik der Fettmigration in Schokoladenprodukten Teil II: Einflu B von Lagertemperatur, Diffusionskoeffizient, Festfettgehalt. Fett-Lipid 98:253–256
19. Ziegleder G, Schwingshandl I (1998) Kinetik der Fettmigration in Schokoladenprodukten Teil III: Fettreif. Fett-Lipid 100:411–415

Chapter 4
Low-Temperature Sweetening in Potato Tubers

4.1 Potato Sprouting

The potato is a staple food in North American and European countries. When exposed to light, potatoes have a tendency to sprout. Sprouting introduces toxic glycoalkaloids (solanine) in the tuber. As well, the potato tuber has a tendency to sweeten during sprouting. Sweetening is a normal biological process associated with sprouting. The sugars released during sprouting serve as an energy source for the growing potato plant. This type of sweetening is called senescent sweetening. Sprouting can be inhibited by the application of sprout inhibitors such as chlorpropham. However, the growing market for "untreated" and "organic" foods necessitates the development of another method of inhibiting sprouting.

One alternative to the use of sprout inhibitors is cold storage. Sprouting can be inhibited by storing potatoes at cold temperatures (2–4 °C). While this method may inhibit sprouting, it does not eliminate sweetening. In fact, cold storage exacerbates sweetening. The storage of potatoes below 8 °C result in an accumulation of sugars. This phenomenon is called "low-temperature sweetening" or "cold-sweetening". It is likely that cold sweetening in plants serve as a cold acclimation survival mechanism. The released sugars depress the freezing point, serving as cryoprotectants by preventing the formation of large ice crystals which may damage plant tissues. As well, plants typically accumulate sugars as a stress-response mechanism in response to conditions such as excess nitrogen in soil, drought and anaerobic conditions in storage (hypoxia).

The presence of the reducing sugars glucose and fructose in sweetened potatoes is undesirable in the processing of such potatoes into fried products such as potato chips and French-cut fries. During frying operations, reducing sugars can react with primary amines on amino acids in what is known as the Maillard reaction. The Maillard reaction results in the formation of dark-colored polymers. Other side reactions such as the Strecker degradation generate flavor compounds which render the product unpalatably bitter. The end result is a fried product that is unaccpetable

A.G. Marangoni, *Kinetic Analysis of Food Systems*,
DOI 10.1007/978-3-319-51292-1_4

to the consumer. The level of reducing sugars in process potatoes must therefore be minimized to ensure the quality of fried potato products.

Despite its apparent dormancy, the potato under storage is a living and metabolically active organism. The energy for respiring potatoes is obtained from the oxidative catabolism of sugars released from starch. Starch in the amyloplast may be broken either hydrolytically (catalyzed by amylases) or phosphorolytically (catalyzed by starch phosphorylase). The products are sugars (dextrins, maltose and glucose) and glucose-1-phosphate, respectively. It is widely held that phosphorylases are the primary starch degrading enzymes in potato tubers during cold sweetening. This is supported by the fact that no increase in the concentration of dextrins (glucose oligosaccharides) and maltose is observed during cold-sweetening. The liberated hexose phosphates are then exported into the cytosol where they are converted into sucrose by sucrose phosphate synthase. Sucrose may then be cleaved by acid invertase to yield glucose and fructose.

The production of sucrose, fructose and glucose in potato tubers is not believed to be controlled by a single factor, but by the interaction of several pathways of carbohydrate metabolism including starch synthesis, glycolysis, mitochondrial respiration and gluconeogenesis. The interactions are complex --- transient oscillatory behavior in sucrose tissue levels has been observed at 4 °C. The exact mechanism responsible for the stress-induced sweetening of potato tissue at the molecular level has not yet been fully elucidated.

Several theories have been proposed to explain the mechanism behind the cold sweetening of potato tubers. Perhaps the earliest theory is that the flow of hexose phosphates (glucose-6-phosphate and fructose-6-phosphate) through the glycolytic Embden-Meyerhof-Parnas pathway is reduced, causing hexoses to accumulate. Low temperatures are believed to inactivate phosphofructokinase-1 (PFK-1) and pyruvate kinase. Both of these enzymes are key regulators of glycolysis in that they catalyze highly exergonic irreversible steps in glycolysis. PFK-1 catalyzes the conversion of fructose-6-phosphate to fructose-1, 6-bisphosphate. This phosphorylation step commits the hexose sugar to glycolysis. Pyruvate kinase catalyzes the conversion of phosphoenolpyruvate to pyruvate with the concomitant synthesis of an ATP molecule.

The theory implies that cold-sweetening is a result of fine metabolic control, that is, cold-sweetening can be observed as soon as a potato is exposed to cold temperatures. Fine metabolic control is an immediate, short-term response to controlling metabolic flux through pathways. It occurs rapidly --- within minutes or seconds. It does not require a significant expenditure of energy. Fine metabolic control involves mechanisms such as covalent modification of enzymes, allosteric control, attachment and dissociation of regulatory subunits, and environmental changes (changes in temperature, pH and substrate concentration).

Some studies, however, support the theory that cold-sweetening is a result of coarse metabolic control. This is supported by the observation of a lag time prior to the onset of cold-sweetening. Coarse metabolic control refers to a gradual long-term response to changing metabolic conditions. It occurs over a period of days to weeks.

It is often accompanied by an expenditure of energy and involves modification of the total amount of enzyme in the cell. It often involves cellular processes such as protein turnover and changes in gene expression.

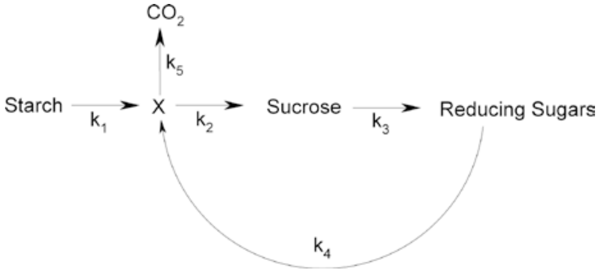
Other theories focus on the state of the cellular and organelle membranes to explain cold sweetening. It is believed that chilling causes physical changes to the physical state of lipid membranes enclosing cellular organelles such as the vacuole (surrounded by a membrane called the tonoplast) and the amyloplast. The vacuole is the cellular organelle responsible for segregating compounds that are toxic to the cell. Cold temperatures alter the permeability of the tonoplast, permitting the leakage of inorganic ions such as phosphate into the cytosol. Chilling also increases the permeability of the amyloplast to phosphate. Phospate is thus translocated into the amyloplast where it inhibits ADP-glucose pyrophosphorylase, a key starch biosynthesis enzyme. This prevents the synthesis of starch at cold temperatures. As well, phosphate stimulates starch phosphorylase, the phosporylase responsible for the breakdown of starch. The overall result is an increase in free sugar.

It has also been observed by researchers that starch was re-synthesized from released sugars when the temperature of cold-stored tubers was increased. Sucrose was formed from starch when potato tubers were transferred from 10 to 2 °C. Starch was re-synthesized when the tubers were moved from 2 to 10 °C. This is the basis of “reconditioning” potatoes. After storage at low temperatures, potatoes are stored at 18–20 °C for several days prior to processing. Reconditioning decreases reducing sugar lwevels, decreasing the likelihood of undesirable browning reactions.

4.2 The Kinetics of Cold Sweetening

A general mathematical model to describe the degradation of starch in mature potato tubers under cold storage was formulated. The model describes the movement of carbon metabolites from starch to sucrose, glucose, fructose and carbon dioxide via the metabolic pool of phosphorylated hexose intermediates ([X]). This is described in Fig. 4.1. The *k* terms denote the rate constants for each step.

Fig. 4.1 A model for the movement of carbon in cold sweetened potatoes



The rate constants in the proposed mechanism can be linked to certain enzyme activities based on knowledge of the factors involved in starch metabolism:

- Starch breakdown - starch phosphorylase, amylase and/or glucosidase activities (k_1)
- Sucrose synthesis - sucrose phosphate synthase activity (k_2)
- Sucrose breakdown - invertase activity and/or vacuolar permeability to sucrose (k_3)
- Entry of metabolites into metabolic pathways - vacuolar permeability to reducing sugars and/or hexokinase activity (k_4)
- Sugar utilization - glycolytic and/or respiratory activities (k_5)

The constants k_1 and k_2 are first-order rate constants while k_3, k_4 and k_5 are zero-order rate constants. The reasons for assuming zero order behaviour in k_3, k_4, k_5 are discussed in the following assumptions:

All effects are purely kinetic in nature. No enzyme induction (the de novo synthesis of new enzyme molecules in response to a signal molecule called the inducer) is involved. It is assumed that LTS involves fine metabolic control rather than coarse metabolic control. This assumption is tenable given that no substantial changes in the V_{max} (indicating no new synthesis of enzyme) of several enzymes involved in carbohydrate metabolism have been observed upon cold storage of potato tubers. In contrast, the Q_{10} of glycolytic enzymes such as ATP-dependent phosphofructokinase and pyruvate kinase increase significantly when the system temperature is increased from 2 to 10 °C.

The Q_{10}, also called the temperature coefficient, is a dimensionless factor by which the reaction rate increases for every 10 °C increment. The mathematical expression for the Q_{10} is given as:

$$Q_{10} = \left(\frac{R_2}{R_1} \right)^{\left(\frac{10}{T_2 - T_1} \right)} \tag{4.1}$$

From this mathematical expression, it can be seen that if $Q_{10} = 1$, then $R_2 = R_1$and that the reaction is not temperature dependent. If $Q_{10} > 1$, then $R_2 > R_1$ and the reaction rate increases with temperature. If $Q_{10} > 1$, then $R_2 < R_1$, indicating that the reaction rate decreases with increasing temperature. The tendency of the Q_{10} to increase from 2 to 10 °C implies that decreasing the temperature from 10 to 2 °C results in a decrease in the Q_{10}, indicating relatively lower reaction rates at cold temperatures.

The assumption that k_3, k_4 and $k_5$are zero-order rate constants holds if the reactions are assumed to be diffusion-controlled, that is, the mass transfer to the catalytic site and not the reaction itself is the rate-limiting step for the reactions listed below. As well, the enzymes responsible for catalyzing the following processes are assumed to be saturated in this particular substrate. The substrate concentration for the given enzymes is assumed to be high enough to permit all of the active sites on the enzyme to be occupied with ligand. The reaction rate under such conditions is assumed to be constant and at the theoretical maximum V_{max}.

- invertase-catalyzed hydrolysis of sucrose to reducing sugars and transfer of sucrose into the vacuole (k_3 is a zero-order rate constant)
- the transfer of reducing sugars out of the vacuole and phosphorylation to hexose phosphates (k_4 is a zero-order rate constant)
- the oxidative respiration of reducing sugars to carbon dioxide (thus k_5 is a zero-order rate constant)

Sucrose hydrolysis by acid invertase will lead to high amounts of fructose and glucose being produced in the vacuole. Removal of these reducing sugars from the vacuole will most probably be mass transfer limited by the capacity of glucose and fructose transporters in the vacuolar membrane. Like enzymes, membrane transporters are saturable in that they display a hyperbolic activity curve relative to the substrate concentration. The rate of removal of glucose and fructose will also depend on how quickly they are phosphorylated into hexose phosphates and utilized in glycolysis and respiration.

It is also assumed that there is no significant starch synthesis occurring in LTS-stressed potato tubers at 2 °C. This assumption is valid considering the tuber under study is a mature tuber, i.e., beyond the growth and development stage. The conditions under which the tuber was stored also inhibits starch synthesis. It has been shown that starch synthesis in potato tubers displays an optimum at about 21.5 °C, with starch synthesis decreasing rapidly as temperature is shifted from this optimum.

The rate constants in this model embody both the activity of enzymes and/or whole enzyme pathways, and/or the transfer of materials in and out of subcellular organelles. In such a phenomenological model, it is not possible to unambiguously differentiate between mass transfer and catalytic kinetic effects. However, it is assumed that the rate constants obtained correspond to enzyme catalytic constants such as k_{cat} since mass transfer limitations should be similar for the two tubers. The differences would therefore be attributed to enzyme activities.

Samples of the potato cultivars Norchip and ND860-2 grown in Cambridge, Ontario, Canada using standard agricultural practices were obtained from Ontario Ministry of Agriculture, Food and Rural Affairs' Agricultural Research Station in Cambridge, Ontario. The Norchip cultivar is sensitive to low temperature sweetening while the ND860-2 cultivar is resistant to cold-sweetening. After curing for 2 weeks at 15 °C, tubers were stored at 2 °C and 95% relative humidity. It must be noted that this model has been shown to represent the data best at 2 °C. At temperatures higher than 2 °C, the data may deviate signficantly from the the predictions made by the model. Random samples of potato tubers were blended and filtered. The sugar content of the supernatant filtrate was quantified using HPLC methods.

The degradation of starch to sucrose, reducing sugars (glucose and fructose) and carbon dioxide depicted in Fig. 4.1 can then be represented by the following set of differential equations:

$$\frac{d[St]}{dt} = -k_1\left([St]-[St_\infty]\right) \tag{4.2}$$

$$\frac{d[X]}{dt} = k_1\left([St]-[St_\infty]\right) - k_2[X] - k_5 + k_4 \tag{4.3}$$

$$\frac{d\left[Suc\right]}{dt} = k_2\left[X\right] - k_3 \tag{4.4}$$

$$\frac{d\left[Red\right]}{dt} = k_3 - k_4 \tag{4.5}$$

A mass balance for the model is given by:

$$\left[St_o\right] + \left[Suc_o\right] + \left[Red_o\right] + \left[CO_{2o}\right] + \left[X_o\right] = \left[St\right] + \left[Suc\right] + \left[Red\right] + \left[CO_2\right] + \left[X\right] \tag{4.6}$$

[St_o] is the initial starch concentration, [St_∞] is the limiting starch concentration as $t \rightarrow \infty$, [Suc_o] is the initial sucrose concentration, [Red_o] is the initial reducing sugar concentration, [CO_{2o}] is the initial carbon dioxide concentration and [X_o] is the initial phosphorylated hexose intermediates concentration. The limiting starch concentration, [St_∞] is the concentration of starch that is unavailable for reaction but is nevertheless quantified during analysis. Initial concentrations of all intermediates were assumed to be zero at $t = 0$.

These equations were simultaneously solved by integration, rearrangement and substitution to obtain the following set of solutions:

$$\left[St\right] = \left[St_o\right] - \left[St_\infty\right]e^{-k_1 t} + St_\infty \tag{4.7}$$

$$\left[X\right] = \frac{k_1\left(\left[St_o\right] - \left[St_\infty\right]\right)}{\left(k_2 - k_1\right)}\left[e^{-k_1 t} - e^{-k_2 t}\right] + \frac{k_4 - k_5}{k_2}\left[1 - e^{-k_2 t}\right] \tag{4.8}$$

$$\begin{aligned}\left[Suc\right] = {} & \frac{k_2\left(\left[St_o\right] - \left[St_\infty\right]\right)}{k_2 - k_1}\left[1 - e^{-k_1 t}\right] + \frac{k_1\left(\left[St_o\right] - \left[St_\infty\right]\right)}{k_2 - k_1}\left[e^{-k_2 t} - 1\right] \\ & + \frac{k_4 - k_5}{k_2^2}\left[e^{-k_2 t} - 1\right] + \left(\frac{k_4 - k_5}{k_2} - k_3\right)t\end{aligned} \tag{4.9}$$

$$\left[Red\right] = \left(k_3 - k_4\right)t \tag{4.10}$$

By difference, from the mass balance, the concentration of CO_2 can be estimated:

$$\left[CO2\right] = \left[St_o\right] + \left[Suc_o\right] + \left[Red_o\right] + \left[CO2_o\right] + \left[X_o\right] - \left[St\right] - \left[Suc\right] - \left[Red\right] - \left[X\right] \tag{4.11}$$

4.3 Relating Kinetic Parameters to Mechanism

Figure 4.2 shows the patterns of sucrose accumulation for ND860-2 and Norchip tubers stored at 2 °C. The solid lines represent the regressed analytic solutions. The tubers grown in 1993 displayed approximately twice the amount of sucrose than the

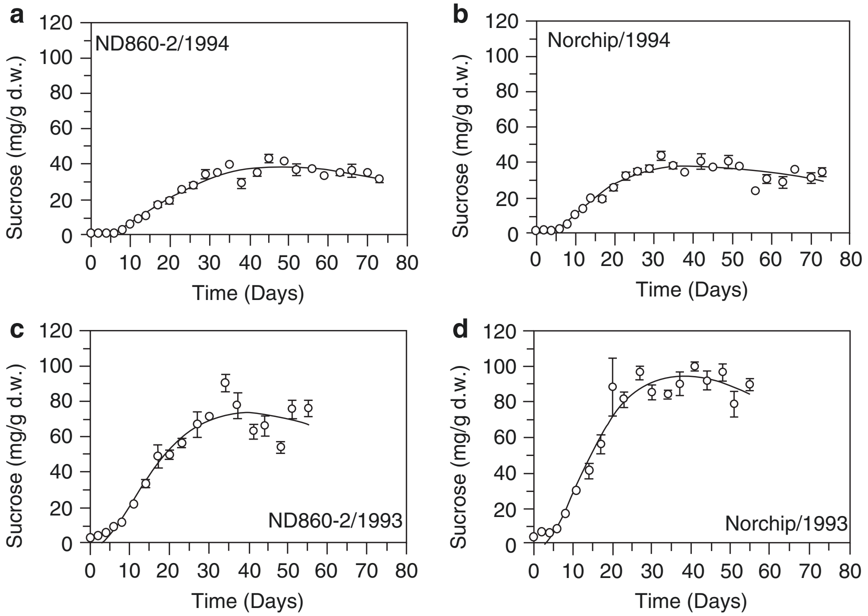


Fig. 4.2 Sucrose accumulation over time of ND860–2 potatoes grown in the year 1994 (**a**) and 1993 (**c**) and Norchip potato tubers grown in the year 1994 (**b**) and 1993 (**d**). Sucrose accumulation was induced by storing the potatoes at 2 °C

tubers grown in 1994. This difference is most likely due to 1994 being an ideal growing year. In contrast, 1993 was a wet and cold year. The environmental stresses in 1993 may account for the accumulation of higher concentrations of sucrose.

Figures 4.3 and 4.4 shows the glucose and fructose concentrations of the respective potato cultivars from the years 1994 and 1993, respectively. The results suggest that the rate and the amount of reducing sugar accumulated depends on the difference between the rate constants for the invertase-catalyzed hydrolysis of sucrose (k_3) and the rate constant for the removal of reducing sugars from the vacuole and/or hexokinase activity (k_4). The rates of accumulation of fructose and glucose were found to be greater for the Norchip (LTS-sensitive) cultivar than for the ND860-2 (LTS-tolerant) cultivar in both years. The rates of glucose and fructose accumulation between the years 1993 and 1994 were not significantly different ($P > 0.05$).

The rate constants and other parameters were derived from fitting the model to the data. These are shown in Table 4.1. Among these, k_1 and k_2 were not significantly different within each cultivar grown in the same season. Furthermore, no significant difference was observed between the k_1 and k_2 of different cultivars grown in the same year. This suggests that the enzymes that catalyze these reactions are not responsible for the difference in sugar accumulation between the two cultivars. However, significant differences in the k_3 (invertase activity) constant were observed between different cultivars grown in the same year. The value of k_3 for Norchip (LTS-sensitive) in 1993 was twice that of ND860-2 (LTS-tolerant). In 1994, however, no significant differences in k_3 between the cultivars were evident ($P > 0.05$).

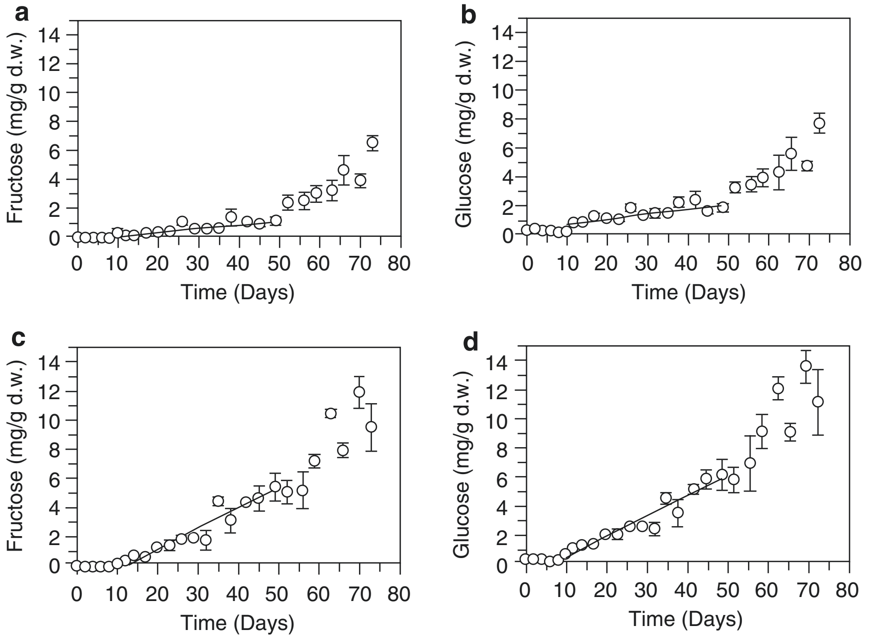


Fig. 4.3 The accumulated level of glucose and fructose in potato tubers grown in 1994 and stored at 2 °C. ND860–2, fructose (**a**); ND860–2, glucose (**b**); Norchip, fructose (**c**); Norchip, glucose (**d**)

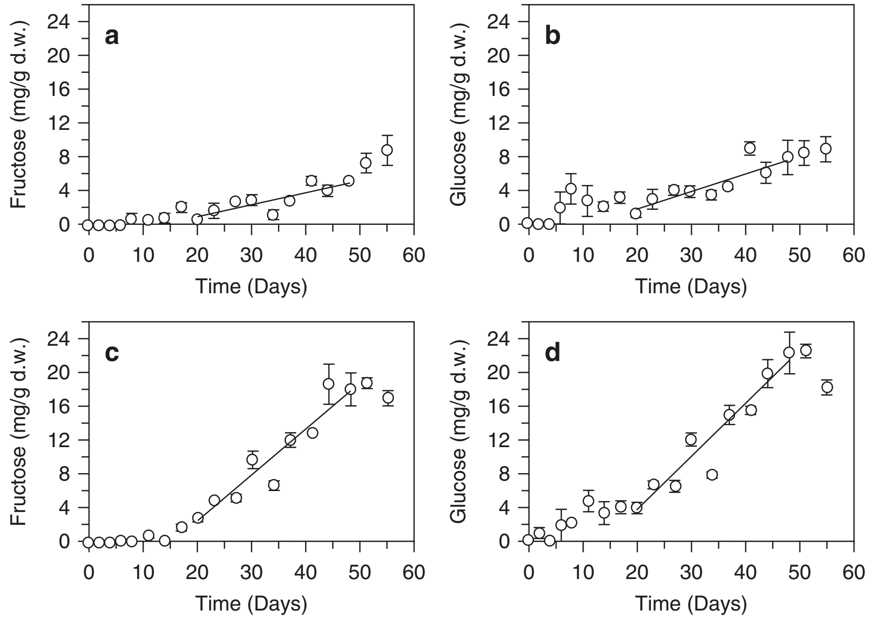


Fig. 4.4 The accumulated level of glucose and fructose in potato tubers grown in 1993 and stored at 2 °C. ND860–2, fructose (**a**); ND860–2, glucose (**b**); Norchip, fructose (**c**); Norchip, glucose (**d**)

Table 4.1 Estimates of kinetic constants as obtained from regression fits to the glucose and fructose content of two potato cultivars grown in the years 1994 and 1993

Parameter	ND860-2 (1993)	ND860-2 (1994)	Norchip (1993)	Norchip (1994)
k_1 (day^{-1})	0.100^{a}	0.073^{a}	0.109^{a}	0.114^{a}
k_2 (day^{-1})	0.098^{a}	0.077^{a}	0.103^{a}	0.108^{a}
k_3 (day^{-1})	1.095^{a}	0.662^{a}	2.282^{b}	1.227^{a}
$k_3 - k_4$ (day^{-1}) Fructose	0.141^{a,e}	0.0217^{b}	0.544^{c}	0.134^{d,e}
$k_3 - k_4$ (day^{-1}) Glucose	0.211^{a}	0.0314^{b}	0.635^{c}	0.136^{d}
$\frac{k_4}{k_5}$	1.007^{a}	1.013^{a}	1.056^{a}	1.095^{a}
$St_0 - St_\infty$ (mg/g dry weight)	126^{a}	65^{b}	160^{c}	77^{b}

Values with the same superscript within a row are not significantly different from each other ($P > 0.05$)

Interestingly, the difference between the rate constant for invertase activity (k_3) and the rate constant for transport of sugars from the vacuoule and/or hexokinase activity (k_4), $k_3 - k_4$ appears to be directly proportional to the amount of sugar that accumulates in the tuber. $k_3 - k_4$ for Norchip for both glucose and fructose was consistently higher than the corresponding value for ND860-2 in both 1993 and 1994. The data suggests that increased invertase activity coupled with an impaired ability to remove sugars from vacuoles in the form of hexose phopshates is responsible for the accumulation of sugar in the system.

The k_4 and k_5 rate constant in both years and cultivars were similar. In fact, the $\frac{k_4}{k_5}$ ratio for all years and cultivars was close to unity. This strongly suggests that k_4 and k_5 describe a similar (or linked) reaction. It is not without reason to suggest that k_4 is the rate constant for the glycolysis pathway (which includes a phosphorylation step) and k_5 describes the remainder of the respiration pathway.

4.4 The Impact of Starch Availability on Cold Sweetening

An interesting result from the fits to the data was the difference observed in the starch available to undergo degradation to hexoses ($[St_o] - [St_\infty]$). The amount of starch degraded was always greater for the Norchip cultivar than for the ND860-2 variety in both years. This result suggests that cold sweetening may be attributable to differences in the availability of starch for LTS-induced degradation. It is possible that the starch from the LTS-sensitive cultivar is more susceptible, or more available, to degradation than starch from LTS-tolerant cultivars. This conclusion would tend to support the theory that starch structure plays a role in LTS-sensitivity or tolerance. A point to consider is that the starch substrate does not become depleted, but rather that the amount of "available starch" initially available becomes depleted. Available starch is the portion of starch that is more susceptible to enzymatic degradation than the main bulk of crystalline starch (estimated to be 6–17% w/w).

4.5 Numerical Simulations

Figure 4.5 shows a simulation of the complete model using the rate constants derived from fitting the experimental data to the analytic solutions of the given differential equations. The simulated patterns agree closely with the experimentally obtained patterns for sucrose and hexose accumulation.

The simulation indicates that there is an initial but temporary increase in phosphorylated hexose intermediates ([*X*]) with a subsequent gradual increase in CO_2 produced. These observations agree with experimental findings published by Isherwood

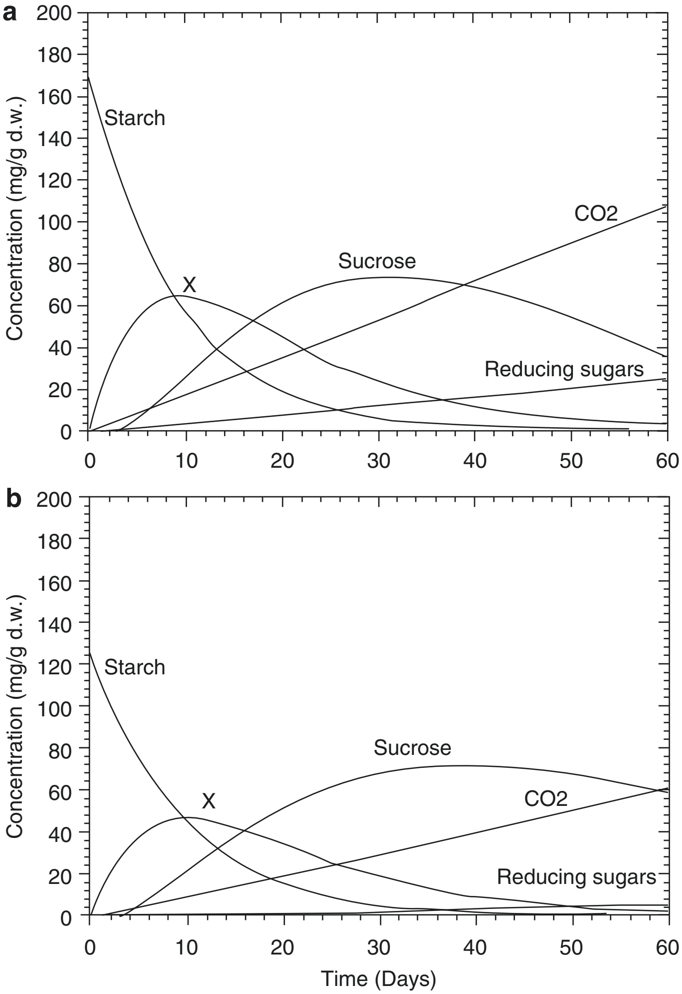


Fig. 4.5 Kinetic simulations of the proposed cold-sweetening mathematical model for the Norchip (**a**) and ND860-2 (**b**) cultivars stored at 2º C. The parameters used were obtained from the non-linear fits of the data to the model as tabulated in Table 4.1

in 1973 and 1976. This prediction is a good "test" of the model, since phosphorylated intermediates were not measured at all in this work. This demonstrates that when kinetic modelling is mechanistic in nature, it can have predictive power, which is more than can be said about statistical models, which mainly predict data within the the data set (interpolation). Power comes from extrapolation!

In conclusion, analysis of sucrose, glucose and fructose accumulation data at 2 °C suggests that low-temperature induced reducing sugar accumulation is due mainly to relative differences between invertase activity and glycolytic/respiratory capacity. A high invertase activity coupled with decreased glycolytic/respiratory capacity results in an increased accumulation of reducing sugars in potato tuber tissue at low temperatures. LTS-sensitive potato cultivars may contain greater amounts of labile starch than tolerant ones.

Bibliography

1. Blenkinsop RW, Copp LJ, Yada RY, Marangoni AG (2003) A proposed role for the anaerobic pathway during low-temperature sweetening in tubers of *Solanum tuberosum*. Physiol Plant 118:206–212
2. Blenkinsop RW, Yada RY, Marangoni AG (2004) Metabolic control of low-temperature sweetening in potato tubers during postharvest storage. Hortic Rev 30:317–354
3. Duplessis PM, Marangoni AG, Yada RY (1996) A mechanism for low temperature induced sugar accumulation in stored potato tubers: the potential rold of the alternative pathway and invertase. Am Potato J 73:483–494
4. Marangoni AG, Duplessis PM, Yada RY (1997) Kinetic model for carbon partitioning in *Solanum tuberosum* tubers stored at 2°C and the mechanism for low temperature stress-induced accumulation of reducing sugars. Biophys Chem 65:211–220
5. Pinhero RG, Copp LJ, Lopez-Amaya C, Marangoni AG, Yada RY (2007) Roles of alcohol dehydrogenase, lactate dehydrogenase and pyruvate decarboxylase in low-temperature sweetening in tolerant and susceptible varieties of potato (*Solanum tuberosum*). Physiol Plant 130:230–239
6. Wismer WV, Marangoni AG, Yada RY (1995) Low-temperature sweetening in roots and tubers. Hortic Rev 17:203–232

Chapter 5
Modelling Sucrose Oscillations in Cold-Sweetened Potato Tubers using a Statistical Approach to Fourier Analysis

5.1 The Fourier Series

The Fourier series is a mathematical tool for expressing a complex sinusoidal (periodic) function as a sum of simpler periodic functions. It is a sub-branch of the mathematical field of harmonic analysis. The Fourier series is named after the French mathematician Joseph Fourier, who utilized Fourier series in his work, "The Analytical Theory of Heat". It is used in several fields of study including image processing (in particular medical imaging), signal analysis, data compression, spectroscopy and acoustics among many others. In the science of foods, it finds an indirect application in the study of heat transfer in foods. In this chapter, sucrose levels in cold-sweetened potato tubers will be modelled using a Fourier series.

5.2 A Periodic Function

A function $f(x)$ is said to be periodic with a period T if

$$f\left(x\right)=f\left(x+nT\right) \tag{5.1}$$

for n = 1, 2,.... The most elementary example of such functions are the sine and cosine functions $\sin(x)$ and $\cos(x)$, each with a period T. The frequency f of a periodic function is the inverse of the period T. The period often has units of time, while frequency has units of time^{-1}.

In lay terms, a function is periodic if it "repeats itself". The period is the "length of time" required for one oscillation. The frequency is the number of oscillations per unit time and is the inverse of the period.

For example, consider a simplified discussion of the phases of the moon. Starting with a "new moon", the moon undergoes a cycle through several phases of the

A.G. Marangoni, *Kinetic Analysis of Food Systems*,
DOI 10.1007/978-3-319-51292-1_5

moon, eventually returning to a `new moon' after 29.5 days. The period of the lunar orbit is 29.5 days. The frequency of this cycle is $\frac{1}{29.5\,days} = 0.0338\ \text{days}^{-1}$ that is 0.0338th of a cycle occurs every day. This phenomenon can be easily modelled using a sinusoidal function.

5.3 The Mathematics of a Fourier Series

Suppose we are given a function f with the real-valued domain $x, f: R \to C$, where $f(x)$ is continuous and periodic with a period T. Also, the definite integral of the function over the interval x_1 to x_2 must be finite. The Fourier series expansion of $f(x)$ is given by:

$$f(t) = \frac{1}{2}a_o + \sum_{n=1}^{\infty}\left[a_n \cos(\omega_n t) + b_n \sin(\omega_n t)\right] \tag{5.2}$$

where n is the harmonic number, ω_n is the angular speed/angular frequency at the nth harmonic, a_n are the even Fourier coefficients and b_n are the odd Fourier coefficients.

The parameters in the Fourier series expansion can be calculated using the following formulas:

$$\omega_n = n\frac{2\pi}{T} \tag{5.3}$$

$$a_n = \frac{2}{T}\int_{x_1}^{x_2} f(x)\cos(\omega_n x)\,dx \tag{5.4}$$

$$b_n = n\frac{2}{T}\int_{x_1}^{x_2} f(x)\sin(\omega_n x)\,dx \tag{5.5}$$

5.4 Modelling Sucrose Levels Oscillation Using a Statistical Approach to Fourier Analysis

After exposure to low temperatures (i.e., <10 °C), potato tubers, as well as many other plants and plant parts, often undergo a phenomenon known as low-temperature sweetening (LTS) which results from the conversion of starch to sugars. Although this phenomenon has been well documented, the causes and mechanisms by which LTS occurs are still not established. Fine metabolic control exerted by allosteric enzymes, coarse metabolic control due to enzyme induction, and mass transfer of reaction substrates and products in and out of subcellular organelles may all play a role in the observed metabolic response to stress.

Excitability, periodic oscillations, polyrythmicity, chaos, bursting or transitions between multiple steady states are possible in metabolic pathways. These dissipative structures are evidence of self-organization in destabilized open systems that are far away from equilibrium and obey nonlinear kinetic laws. Such non-equilibrium conditions prevail in cells, tissues and whole organisms, which are highly compartmentalized structures where diffusion is coupled with biochemical reactions. Oscillations can therefore occur at all levels where regulation is exerted.

A classic example of a metabolic oscillation is that of glycolysis. Glycolytic oscillations have been studied in yeast and muscle extracts and intact yeast cells. They are solely attributed to the positive feedback regulation of 6-phosphofructokinase (an allosteric enzyme) by the reaction product adenosine diphosphate (ADP) and adenosine monophosphate (AMP). AMP is linked to ADP by the activity of adenylate cyclase.

While the physiological significance of glycolytic oscillations remains unclear, excitability and oscillations in cyclic nucleotides and calcium concentrations are an integral part of the intercellular communication system of the cellular slime mould *Dictyostelium discoideum*. Cyclic nucleotide waves, induced by starvation, are generated and relayed by the individual amoebae, acting as a chemotactic attractant. These chemical waves regulate the aggregation of over 10^5 individual amoebae into a motile multicellular organism which eventually transforms into a fruiting body with differentiated stalk and spore cells.

The glycolytic oscillator is an example of metabolic oscillations in unicellular organisms and in tissue extracts while the cyclic nucleotide signalling system of *Dictyostelium discoideum* is an example of a metabolic oscillator in a simple multicellular organism. This work presents evidence of a stress-induced transient oscillation in the sucrose metabolism of intact potato tubers. The transient oscillation is an example of a non-monotone approach of a complex biological system to a new steady state.

Solanum tuberosum cv. Norchip (LTS-sensitive) and ND860–2 (LTS-tolerant) plants were grown from seed during the summers of 1993 and 1994 at the Cambridge Agricultural Research Station, Ontario Ministry of Agriculture, Food and Rural Affairs (Cambridge, ON) using standard agronomical practices. Tubers were harvested manually in late September and stored for 2 weeks at 15 °C prior to storage at 12, 10, 8, 4 and 2 °C (1993), and 16, 12, 8, 4, and 2 °C (1994). Tubers were not sprout inhibited in either growing season. Sucrose, glucose and fructose concentrations (mg/g dry weight) were determined by high performance liquid chromatography (HPLC) as described by Wilson and co-workers.

Sucrose accumulation patterns were modelled using Fourier series analysis (5). The general form of the Fourier series used was:

$$[Sucrose]\left(mg\,g^{-1}\,d.w.\right) = A_0 + \sum_{k=1}^{n} A_i \sin\left(k\omega t + \varphi_k\right) \tag{5.6}$$

where A_o is average initial sucrose concentration, t is the time in days, A_k corresponds to the amplitude of the harmonic of frequency $k\omega$, where k is an integer ranging from 1 to n, while φ_k corresponds to the phase angle of the particular harmonic. The Fourier series were fitted to the experimental data using nonlinear least squares methods with robust data weighting to exclude outliers. The software program used the standard Marquart-Levenburg method using a numerical second order method to calculate partial differentials. The criterion for convergence was a less than 0.01% change in the reduced chi-square (χ^2) value upon variation of parameters. Robust data weighting was performed using the method of Mosteller and Tukey, as implemented by Duggelby.

The experimental points for the LTS-sensitive Norchip tubers grown in 1993 and stored at 4 °C were initially fitted to a Fourier series with a period equal to the full data length T ($k = 1$). Additional, higher order harmonics were then attempted by increasing sequentially the value for k, where period components are given by T/k. Each subsequent harmonic added to the $k = 1$ series was evaluated by testing for significance of improvements ($p < 0.05$) in the reduced chi-square value (χ^2) for the fit (F-test). The F-test was performed by comparing the ratios of reduced chi-square values. Nonlinear fits using simple weighting and robust weighting were identical; hence, none of the data points were considered outliers in our analysis. A higher order Fourier series ($k > 1$) was adopted only when the fit to the data was significantly improved over its $k = 1$ counterpart ($p < 0.05$). The process was then repeated using the new higher order Fourier series as the starting point for comparisons. Each subsequent harmonic added to the series was evaluated by testing for significance of improvements ($p < 0.05$) in the reduced chi-square values (χ^2) for the fit (F-test) as described before.

A seven harmonic series fitted the data significantly better than a one harmonic series ($p < 0.05$) in 1993. No significant improvements in the fit to the data ($p > 0.05$) were observed by adding a greater number of harmonics. The one-harmonic fit to the data was used to detrend this non-stationary time series.

For 1994, a seven harmonic series also fitted the data significantly better than a one harmonic series ($p < 0.05$). No significant improvements in the fit to the data ($p > 0.05$) were observed by adding a greater number of harmonics. For the 1994 data, the detrending procedure was slightly different. A linear fit (using robust data weighting) through the data, from day 6 onwards, was subtracted from the experimental points resulting in a stationary time series. The aim of the detrending procedures was to produce stationary time series.

The 1993 and 1994 LTS-tolerant cultivar (ND860–2) data at 4 °C were transformed by simply subtracting the average steady-state value of sucrose from each data point, usually the average of all values after day 2. Its dynamic behaviour was not oscillatory in nature.

Sucrose concentrations of potato tubers stored at 10 °C or higher maintained a steady-state value throughout the sampling period (data not shown). However, when potato tubers are stored below 10 °C, sucrose concentrations increase to a new steady state. This was clearly shown for the 1993 and 1994 tubers (Fig. 5.1). This discontinuity at 10 °C may constitute a bifurcation point in sucrose metabolism, where exposure to temperatures below this critical value cause the system, i.e. the

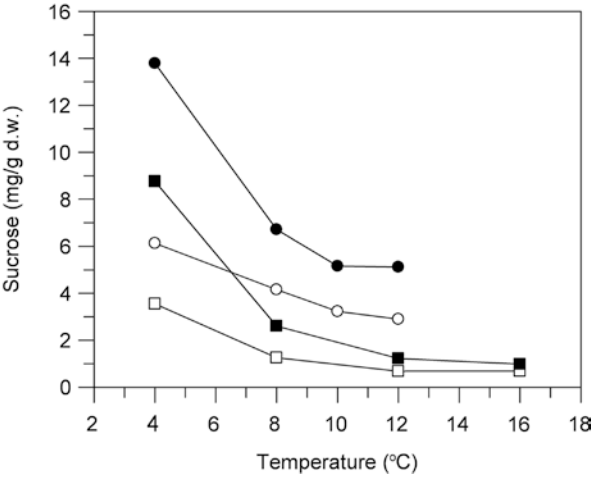


Fig. 5.1 Sucrose tuber tissue concentrations (mg/g d.w.) as a function of temperature for low-temperature sweetening (LTS) sensitive *Solanum tuberosum* Norchip tubers grown in 1993 (●) and 1994 (□), and LTS-tolerant *Solanum tuberosum* ND860–2 tubers grown in 1993 (○) and 1994 (□). Below 10 °C, sucrose tissue concentrations increase significantly ($p < 0.05$) with decreasing temperatures

potato tuber, to jump into a different state that produces higher sucrose concentration. The simplest case of a bifurcation is the transition from one steady state to another. More complex cases include transitions from a steady state to oscillations, or from oscillations to chaos.

At a critical temperature change from the apparent bifurcation point at 10 °C (ΔT), complex dynamic behaviour in sucrose metabolism of potato tubers was observed. This critical temperature distance depends on the relative tolerance or sensitivity of the tissue to low temperature stress. For a LTS-sensitive cultivar (Norchip), ΔT is 6 °C. The average sucrose concentration values at 4 °C for the LTS-sensitive cultivar shown in Fig. 5.1 are averages of the fluctuating values observed in tissue sucrose levels at this temperature (Fig. 5.2a). A true steady-state concentration was not achieved at this temperature.

Figure 5.2a shows the complex pattern of sucrose accumulation in 1993-grown Norchip tubers during a 55-day period at 4 °C, while Fig. 5.3a shows the complex pattern of sucrose accumulation in 1994-grown Norchip tubers during a 73-day period at 4 °C. Sucrose concentration did not increase monotonically from one state to another, but rather transiently oscillated towards the new state. The non-monotone behaviour observed in these 2 years, for potatoes grown under vastly different environmental field conditions, was strikingly similar. For comparison purposes, the time axis of both patterns were normalized to a relative time. This parameter is defined as the actual storage time divided by the time required to reach the apex of the first large oscillation (27 days in 1993 and 49 days in 1994). When the 1993 and 1994 data were compared in terms of their relative times (Fig. 5.4), striking pattern similarities were evident; however, the period of the oscillations observed in 1993 was shorter than that observed in 1994. Differences in growing conditions (temperature, precipitation, environmental stresses) between the 2 years may have influenced initial tuber carbohydrate metabolism.

For proper mathematical analysis, time series have to be detrended; that is, the underlying tendency of the series must be removed from their oscillatory behaviour

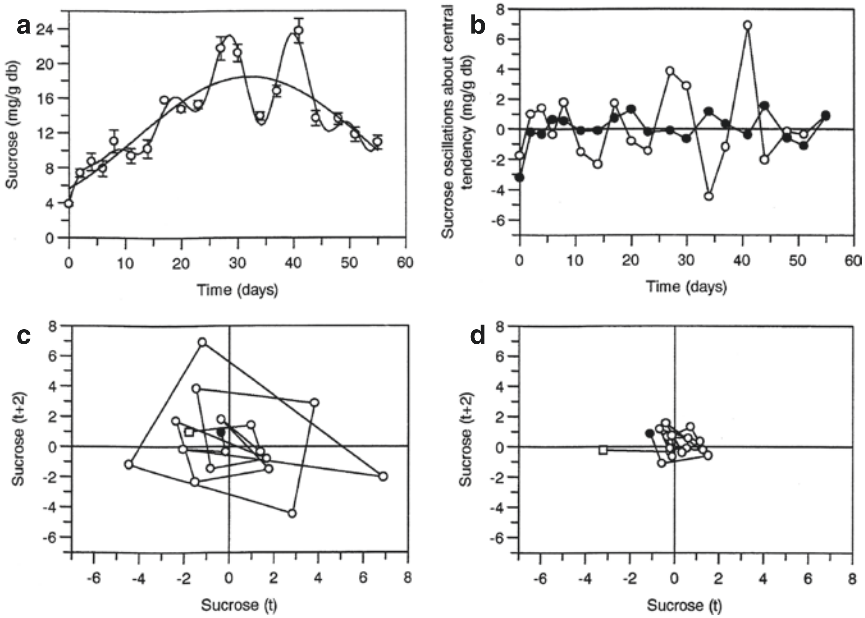


Fig. 5.2 Tubers grown in 1993. (**a**) Oscillatory pattern of sucrose accumulation (mg/g d.w.) in potato tuber tissue from the low-temperature sweetening (*LTS*) sensitive Norchip cultivar as a function of storage time at 4 °C. Each point in time represents the average of three separate experiments and their standard errors. (**b**) Stationary (detrended) time series for sucrose tuber tissue concentrations in LTS sensitive (○) and LTS tolerant (●) cultivars. (**c**) Transition function portrait for the LTS-sensitive cultivar tuber sucrose concentrations. The sensitive cultivar's sucrose levels oscillate at 4 °C because, unlike the tolerant tuber, it is not able to maintain a steady-state tissue concentration. (**d**) Transition function portrait for the LTS-tolerant cultivar tuber sucrose concentrations. The tolerant cultivar manages to keep sucrose levels within a narrow margin of concentrations at 4 °C. In both transition function portraits, (□) and (●) represent the beginning and end of the time series, respectively

(5). These non-stationary time series were detrended as outlined before. The resulting stationary time series are shown in Fig. 5.2b (1993) and Fig. 5.3b (1994). Our results show that Norchip tuber sucrose levels transiently oscillate about a central accumulation tendency towards the new steady state. In contrast, for a LTS-tolerant (ND 860–2) tuber, the transient oscillations are much smaller, or nonexistent, and sucrose levels increase monotonically to a new steady state. This effect becomes quite obvious in the transition function portraits (iterated map) for both time series (Fig. 5.2c and 5.2d for 1993 and Fig. 5.3c and 5.3d for 1994). The sensitive cultivar loses part of its metabolic control and the tissue sucrose levels transiently oscillate (Figs. 5.2c and 5.3c), while the tolerant cultivar manages to keep a tight metabolic control over its sucrose tissue levels (Figs. 5.2d and 5.3d).

At 2 °C, tissue sucrose concentrations increased in a sigmoidal, non-oscillatory fashion (data not shown) to an average of 89 mg/g d.w. (Norchip) and 66 mg/g d.w. (ND 860–2) in 1993 and 35 mg/g d.w. (Norchip) and 35 mg/g d.w. (ND860–2) in 1994. This situation possibly represents a complete loss of metabolic control. We

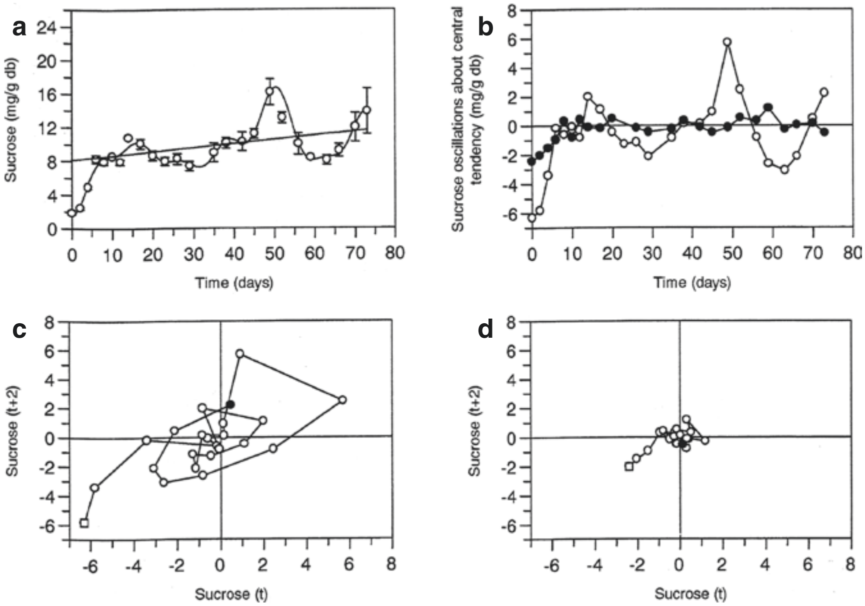


Fig. 5.3 Tubers grown in 1994. (**a**) Oscillatory pattern of sucrose accumulation (mg/g on a dry basis) in potato tuber tissue from the low-temperature sweetening (*LTS*) sensitive Norchip cultivar as a function of storage time at 4 °C. Each point in time represents the average of three separate experiments and their standard errors. (**b**) Stationary (detrended) time series for sucrose tuber tissue concentrations in LTS sensitive (○) and LTS tolerant (●) cultivars. (**c**) Transition function portrait for the LTS-sensitive cultivar tuber sucrose concentrations. The sensitive cultivar's sucrose levels oscillate at 4 °C because, unlike the tolerant tuber, it is not able to maintain a steady-state tissue concentration. (**d**) Transition function portrait for the LTS-tolerant cultivar tuber sucrose concentrations. The tolerant cultivar manages to keep sucrose levels within a narrow margin of concentrations at 4 °C. In both transition function portraits, (□) and (●) represent the beginning and end of the time series, respectively

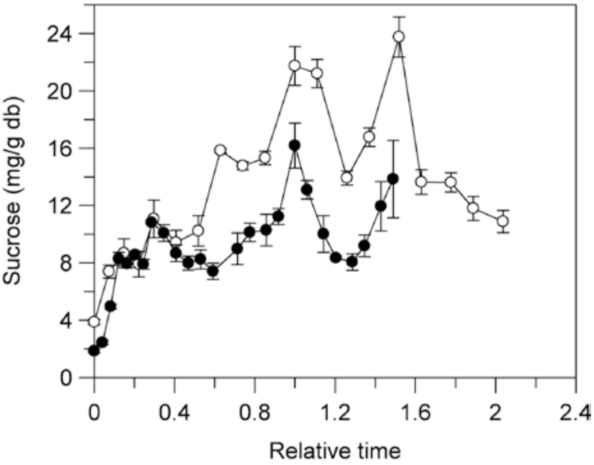


Fig. 5.4 Oscillations in sucrose tuber tissue concentrations in the low-temperature sweetening (*LTS*) sensitive Norchip cultivars at 4 °C for 1993 (○) and 1994 (●) grown tubers. The relative time corresponds to the time divided by the time required to reach the apex of the first oscillation

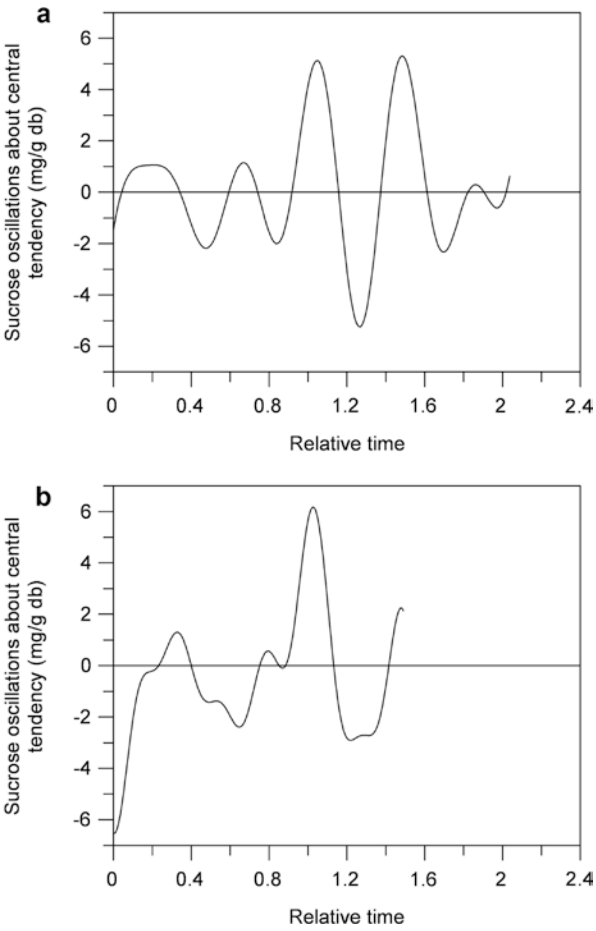


Fig. 5.5 Oscillations in sucrose tuber tissue concentrations in the low-temperature sweetening sensitive Norchip cultivar at 4 °C for 1993 (**a**) and 1994 (**b**) grown tubers. Data points were generated every 0.25 days using Fourier series analysis and series data detrending as described in the text. The relative time corresponds to the time divided by the time required to reach the apex of the first oscillation

therefore postulate that at 4 °C, the LTS-sensitive tuber's carbohydrate metabolism becomes temporarily deregulated, undergoes oscillations but manages to adapt to the decreased temperature and reach a new steady state. In contrast, at 2 °C all metabolic regulation or capacity for recovery are lost.

The sucrose oscillations were analyzed using Fourier series analysis in order to generate a preliminary power spectrum. The power is defined as the square of the amplitude A_k at the corresponding harmonic $k\omega$. The patterns shown in Fig. 5.5 were generated by subtracting the values of the seven harmonic Fourier series from the one-harmonic Fourier series fit to the experimental data every 0.25 days. For the 1993 data (Fig. 5.5a), the resulting stationary time series was best modelled by a one-harmonic Fourier series (a sine wave) with a period of 11.5 days (0.55 radians/day), i.e., the time series displayed only one predominant frequency. No improvements in the fit to the data were observed when higher harmonics were added ($p > 0.05$). For the 1994 data (Fig. 5.5b), the resulting stationary time series was best modelled by a nine-harmonic Fourier series with a period of 80 days (0.078 radians/day). Each subsequent addition of a harmonic was significant ($p < 0.05$), however,

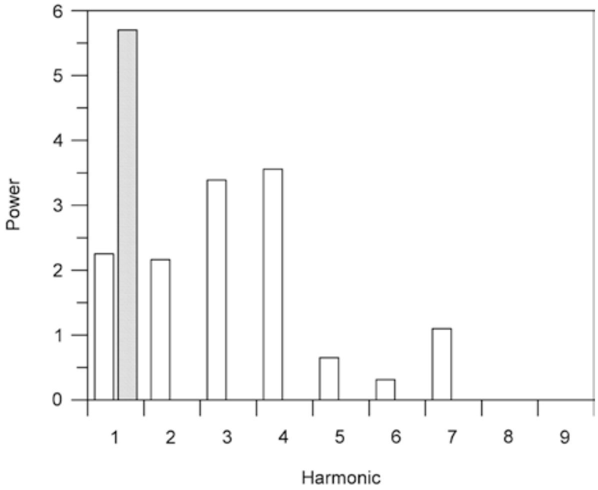


Fig. 5.6 Power spectra for the 1993 (■) and 1994 (□) sucrose oscillations observed in the low-temperature sweetening sensitive Norchip potato tubers

the contribution was becoming smaller with each new harmonic added, and therefore the process was stopped at the ninth harmonic.

A power spectrum (square of the amplitude of the harmonic vs. harmonic number) was generated for these waves (Fig. 5.6). For the 1993 data, the power spectrum could only be generated up to three harmonic components because of constraints imposed by the Nyquist theorem: "The maximum frequency (minimum period) that can be estimated is one corresponding to half (twice) the sampling frequency (period)". Our minimum sampling period was 2 days. For the 1994 data, we could have generated a spectrum of up to 20 harmonic components. We doubled the requirements of the Nyquist theorem by estimating only up to 10 harmonics for the Norchip cultivar's oscillations. No aliasing was observed in any of the analysis. We did not have enough data for this, but a plot of the log of the square of the amplitude (the power) of a particular harmonic vs. the log of the harmonic number would yield a slope, *f*. The value of this slope has been associated with specific process, for example, if the thee slope of this log-log plot is:

0 – White noise ($1/f^0$ noise)
1 – Mandelbrot noise or chaos (1/f noise)
2 – Brownian motion or diffusion process ($1/f^2$ noise)

Dynamic processes (time histories of biological variables) can be generated by deterministic processes or by random processes. Deterministic processes: (cause-effect) can be periodic, quasi-periodic or chaotic. Power spectra of these processes have patterns. They can be broad and have a 1/f pattern (chaotic) or be narrow and display only a single frequency component (periodic oscillations). Random processes: random, probabilistic in nature. Power spectra are broad and featureless and are due to white noise.

This analysis clearly demonstrated the presence of only one harmonic in the periodic motion of the 1993 sucrose oscillations. Had the data not been detrended, a power spectrum could have been erroneously generated using the seven harmonic series, or even worse, assumed to be an infinite number of harmonics as in classic

Fourier analysis. For the 1994 data, a more complex oscillation was observed with several frequency components. Even though the oscillations observed in 1993 and 1994 were quite similar when expressed in terms of relative time, subtle differences in the oscillatory patterns were evident. Whether the 1994 oscillation contains random noise or is deterministic in nature is unclear at the moment, however, the pattern observed in Fig. 5.6 is reminiscent of white noise.

Glucose and fructose are the products of sucrose degradation and also displayed complex dynamic behaviour in the sensitive cultivar. However, the observed patterns were not periodic and their relationship to the sucrose oscillations was not evident. As well, the signal to noise ratio is smaller for fructose and glucose relative to that of sucrose and it becomes difficult to discern between a true oscillations and random noise.

This chapter highlights a statistical approach to Fourier analysis. It is my view that this approach is better suited to the study of complex systems where the usual assumption that an infinite number of frequencies can be present in a time series is probably not correct. As opposed to traditional Fourier Analysis, no assumptions of data orthogonality ($\Sigma xy = 0$) are made and we do not require equally spaced data points. Our approach does make the assumption that errors are random, that is, they are normally distributed with mean 0 and standard deviation sdx. Here we fitted the Fourier function to the experimental time series, starting with a sine wave ($k = 1$), and added one harmonic at a time while testing for an improvement in the fit of the series to the data. We only added a higher harmonic if the fit using the higher order harmonic fitted the data statistically significantly better than the lower harmonic series. We stopped when further improvement was not observed. This allowed us to characterize the complex transient behaviour in the sugar metabolism of a whole living system (the potato). This approach could be used in many other biological and food systems.

As a final philosophical comment to the fact that oscillations and instabilities are observed in biological systems, we paraphrase the words of Yates:

"The very nonlinear world of chaotic dynamics and fractal geometries seems to me to justify the term homeodynamics for homeostasis. In the stability of biological systems over space and time, homeostasis is linear and limited while homeodynamics is nonlinear and general. Homeodynamics carries with it the potential for a deeper understanding of what it means for a complex system to be stable and yet show very rich behaviors, such as chaos and fractals, while including the development of individuals and evolution of species."

5.5 A "How-To" Analytical Toolbox for the Study of Instabilities and Oscillations in Complex Systems

1. Collect your data as a string of numbers (lots), usually equally spaced in time. Be careful to choose the proper sampling intervals - use common sense. Do not smooth or filter any data. Be sure to collect at least 200 points in time.
2. Remove any non-stationarity in your time series before any analysis is carried out. You may choose to analyze a shorter record length, however, this is not good practice.

3. Perform spectral analysis and examine the power spectrum for evidence of periodicities and broad-band, background noise (1/1, 1/f, $1/f^2$). Fit your time series to a one-harmonic Fourier Series (FS), obtain a chi-square value. Fit your harmonic to a second order FS and determine the chi-square value. Test for significance of improvement in the fit (F test). Continue until you find significant improvement. Start the process over. Do not fit to a higher number of harmonics unless you observe significant improvements in the fit. Once you obtain your FS, square each amplitude component and plot a power spectrum.
4. Plot Y(t + 1) vs. Y(t) along the whole time series. The resulting iterated map gives the Transition Function. If the process is random, a random scatter of points will be observed. If the process is oscillatory in nature and stable, a limit cycle will be observed. If the process is a dampened oscillation, an inward spiral will be observed. If the process is chaotic, a strange attractor may be evident.
5. Calculate a Correlation Dimension or Apparent Entropy. The nature of dynamic processes may be classified according to the values obtained. Most of these dimensionality calculations are "souped-up" serial correlation coefficients:

$$r_i = \Sigma y(t)\,y(t+i))/\left(\Sigma y(t)^2\,\Sigma y(t+i)^2\right)^{1/2} \quad \text{where}\, i = 1,2,3,4,\ldots\ldots,n$$

With a serial correlation coefficient you can generate a lagk function.

Bibliography

1. Berridge MJ, Rapp PE (1979) A comparative survey of the function, mechanism and control of cellular oscillators. J Exp Biol 81:217–279
2. Blenkinsop RW, Yada RY, Marangoni AG (2004) Metabolic control of low-temperature sweetening in potato tubers during postharvest storage. Hortic Rev 30:317–354
3. Duggleby RG (1981) A nonlinear regression program for small computers. Anal Biochem 110:9–18
4. Goldbeter A (1997) Biochemical oscillations and cellular rhythms. The molecular bases of periodic and chaotic behaviour. Cambridge University Press, Cambridge, MA
5. Hess B (1971) The glycolytic oscillator. J Exp Biol 81:7–14
6. Hess B, Markus M (1987) Order and chaos in biochemistry. Trends Biochem Sci 12:45–48
7. Johnson ML, Brand L (eds) (1994) Numerical computer methods, part B. Methods in enzymology, vol 240. Academic Press, New York
8. Marangoni AG, Duplessis PM, Lencki RW, Yada RY (1996) Low-temperature stress induces transient oscillations in sucrose metabolism in *Solanum tuberosum*. Biophys Chem 61:177–184
9. Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters. J Soc Ind Appl Math 11(2):431–441
10. Rapp PE (1971) An atlas of cellular oscillators. J Exp Biol 81:281–306
11. Sel'Kov EE (1968) Self-oscillations in glycolysis 1. A simple kinetic model. Eur J Biochem 4:79–86
12. Wismer WV, Marangoni AG, Yada RY (1995) Low-temperature sweetening in roots and tubers. Hortic Rev 17:203–232
13. Yates FE (1992) Fractal applications in biology: scaling time in biochemical networks. Methods Enzymol 210:636–675

Chapter 6
Biodiesel Synthesis via Transesterification of Soybean Oil with Methanol

6.1 The Need for Alternative Fuel Sources

Given the rapid rate of consumption of the world’s mineral oil resources, there is a great likelihood of an energy crisis in the future. Given the vital role that oil plays as an energy source in industrialized socieities, an alternative energy source is necessary to ensure the continuation of civilization. Fuels derived from biological sources, or biofuels, are being explored for use as an alternative energy source. A wide variety of organic compounds are being explored for use as biofuels. Chief among these are ethanol and butanol. Ethanol can be produced by the ethanogenic fermentation of carbohydrates by yeasts (Saccharomyces) and the bacteria *Zymomonas mobilis*. Butanol is produced by the Weizman process wherein sugar substrates are fermented by anaerobic *Clostridium acetobutylicum* to produce butanol and acetone.

6.2 Biodiesel

The idea of using biomaterials as a fuel source is not a recent idea. Over a century ago, Rudolf Diesel, the inventor of the diesel engine, tested vegetable oil as a fuel for his engine. Biological materials used as fuel for diesel engines is called biodiesel. The introduction of cheap petroleum sources eclipsed the use of vegetable oils as a fuel source. Ever since, vegetable oils have only been used during fuel shortages. It must also be noted that biodiesel need not necessarily be produced from vegetable oil sources alone. Animal fats, single-cell oils (oils from microorganisms), terpenes and latexes can also be used to produce biodiesel. However, for the remainder of this discussion, the term “biodiesel” shall refer specifically to biodiesel made from vegetable oil sources. Biodiesel need not be made from high-quality virgin oil. Biodiesel made from fresh vegetable oil is relatively expensive. Used cooking oils can be used as the reaction substrate although the high free fatty acid and moisture content of

A.G. Marangoni, *Kinetic Analysis of Food Systems*,
DOI 10.1007/978-3-319-51292-1_6

such oils present technological challenges that must be first overcome before the use of waste frying oil to produce biodiesel can be realized.

Biodiesel has several advantages over conventional diesel fuel or petrodiesel. First and foremost, it is a renewable energy source. It can be made from readily available raw materials which can be agriculturally grown. It has the image of being more environmentally friendly (low emission profiles). The heat content of biodiesel is also similar to that of petrodiesel. The long chain fatty acid moeity of biodiesel is chemically similar to the long-chain hydrocarbons typical of petrodiesel. Biodiesel contains approximately 80% of the heat content of diesel fuel from petroleum sources. Other advantages include non-flammability and biodegradability. Existing engines can also be made compatible with biodiesel without having to undergo major modifications.

Several challenges need to be overcome before triglyceride-based fuels can be economically used. Vegetable oils can be very viscous, particularly if they are not degummed. Their high viscosity causes them to block fuel injector lines in diesel engines. This can result in the build-up of carbon deposits (coking) during combustion. High viscosity along with low volatility makes it difficult to introduce the fuel into the combustion chamber as a fine mist, especially in cold weather. Engine lubricating oils contaminated by vegetable oil can also thicken and gel, requiring frequent replacement of engine lubricant. Polyunsaturated fatty acids can also polymerize and gel at high operating temperatures or during prolonged storage. These problems are often manifest only when the engine has been running on vegetable oils for an extended period of time. In many cases, running diesel engines on biodiesel for short periods of time do not lead to any adverse effects.

6.3 Biodiesel Production

6.3.1 Blending

The simplest and most crude form of biodiesel is a mixture of vegetable oils and conventional diesel fuel. The use of such mixtures in diesel engines was pioneered by the Caterpillar company in the 1980's in cooperation with the Brazilian government. Research carried out on triglyceride-diesel blends found that blends containing a high proportion of unmodified vegetable oils were simply not practical. 20% vegetable oil + 80% diesel fuel was successfully used as a fuel and some experiments were able to satisfactorily use a 50:50 ratio. Despite this, unmodified vegetable oils presented problems in the form of an increase in viscosity as a higher proporation of vegetable oils was used.

6.3.2 Derivatization

The viscosity increase can be ameliorated by derivatizing the vegetable oil to make it more compatible with existing engines. Most of these derivatization processes result in the chemical transformation of the oil but some processes such as microemulsification, do not result in chemical changes. Biodiesel microemulsions of vegetable oil can be formed with organic solvents such as methanol, ethanol and 1-butanol. Microemulsions are colloidal dispersions of fluid microstructures with diameters ranging from 10 to 100 nm.

Microemulsions are formed spontaneously by mixing two normally immiscible fluids along with a surfactant. Microemulsions, unlike ordinary emulsions, do not need to be sheared to be formed. Microemulsification was found to lower the viscosity, improving the spray characteristics and injection of fuel into the combustion chamber.

6.3.3 Pyrolysis

Biodiesel can also be produced from the process of pyrolysis or thermal cracking. Pyrolysis is the conversion of one substance into another by means of heat, with or without a catalyst. Pyrolysis is different from combustion in that it is conducted in the absence of oxygen while oxygen is a necessary prerequisite for combustion. In the pyrolytic production of biodiesel (or rather conventional diesel from biological sources), oils are pyrolyzed to form hydrocarbons from vegetable oils. The reaction usually yields shorter chain hydrocarbons characteristic of gasoline (octane) instead of long chain hydrocarbons that comprise diesel. Pyrolysis often results in the formation of a wide variety of products, making it difficult to characterize the reaction throroughly. Research on the thermal cracking of several vegetable oils concluded that it was possible to produce hydrocarbon-rich fractions. However, the capital investment for pyrolytic equipment is relatively high even for moderate outputs. Fuels made from pyrolytic cracking contain a high amount of un-oxygenated hydrocarbons. As well, fuels made via pyrolytic cracking have a relatively high particulate emission profile. Fuels made from pyrolytic reactions are non-oxygenated, that is, they do not contain any oxygen in their chemical structures. This is a consequence of performing the pyrolytic reaction at anaerobic conditions. The oxygen content of a fuel influences its particulate emission profile. An oxygenated fuel is believed to result in a more complete combustion than non-oxygenated fuels. Incomplete combustion of fuel (usually stemming from a lack of oxygen during combustion) leads to the formation of particulates (soot) and carbon monoxide. These are major air pollutants. Thus, fuels made from pyrolytic cracking are viewed as less environmentally friendly than the oxygenated biodiesel. A fuel can be oxygenated by adding combustible oxygen-containing additives. Petrodiesel is, in fact, commonly oxygenated by adding alcohols (ethanol, methanol and butanol) or ethers (methyl tert-butyl ether and ethyl tertiary butyl ether) to give a “cleaner” fuel.

6.3.4 Transesterification

The most commonly employed method of biodiesel production is transesterification. It is currently the process of choice for producing biodiesel from vegetable oils. Transesterification, or more specifically alcoholysis, or methanolysis if methanol is

Fig. 6.1 A reaction for the transesterification of a triacylglycerol to yield a fatty acid methyl ester (*FAME*) and diacylglycerol

used, is the reaction by which the fatty acid moiety of a triglyceride is removed and esterified to an alcohol to produce fatty acid esters and glycerol. The reaction is conducted in the presence of a catalyst, which may be a chemical or an enzyme. This reaction is shown in Fig. 6.1. The reaction mechanism by which Fatty Acid Methyl Esters (FAMEs) products are formed is called the carbonyl addition scheme. In this scheme, the negative oxygen moeity of methoxide (or a similar catalyst) conducts a nucleophilic attack on the carbonyl carbon of a triglyceride to form a tetrahedral intermediate. This reaction intermediate decomposes to yield a glycerylate ion and a FAME. The glycerylate ion reacts with methanol or any alcohol to regenerate the basic catalyst.

Biodiesel made of these fatty acid esters is non-toxic and biodegradable. This type of fuel also does not contain sulfur, reducing pollution from combustion processes. Moreover, the glycerol produced can be recovered and sold to defray the high cost of tranesterification. Methanol is often used as the alcohol of choice, due to its low cost and chemical advantages. It is the most polar primary alcohol. Basic catalysts such as NaOH are easily dissolved in methanol. Ethanol can also be used. The most commonly studied vegetable oils for their potential to be used as biodiesel stocks are soybean oil, sunflower oil and rapeseed oil. These are the most commonly produced vegetable oils in countries pioneering the use of biodiesel. The oils that can be used are by no means limited to these. Transesterification produces a mixture of esters, mono- and diglycerides (resulting from incomplete reaction), alcohol, glycerol and catalyst. This reaction end product mixture must be purified to obtain mostly fatty acid esters, the primary component of biodiesel. Purification of the biodiesel can be carried out using a variety of methods ranging from simple distillation to ion exchange chromatography.

Chemical transesterification is usually carried out in the presence of a basic or acid catalyst with an excess of alcohol. The excess of alcohol is necessary since the transesterification reaction is reversible. Excess alcohol (>3:1 alcohol:triglyceride) ensures that the reverse reaction is limited compared to the forward reaction. A high molar ratio of alcohol to triglycerides is necessary to ensure a high reaction yield.

Basic catalysts are preferred over acid catalysts in the commercial production of biodiesel. Acid catalysis requires a higher molar ratio of alcohol to triglyceride, usually in the order of 15:1–30:1. Basic catalysts only require molar ratios of 3:1–6:1. Commonly-used basic catalysts include KOH, NaOH, carbonates and alkoxides such as methoxide and ethoxide. Note that methoxide and ethoxide are generated when the hydroxyl group in the corresponding alcohol (methanol or ethanol) is deprotonated. This deprotonation reaction inevitably results in the formation of minute amounts of moisture, which may hinder later stages of the reaction. Because of this, it is preferable to use straight methoxide or ethoxide as direct addition of these catalysts (in powder form) do not generate moisture. However, alkoxides are relatively expensive compared to strong bases. Large-scale transesterification typically uses strong base.

Reactions catalyzed with basic catalysts can proceed at a rate 4000 times greater than that of reactions with acid catalysts. However, several challenges need to be overcome in order to use basic catalysts efficiently. Basic catalysts are very sensitive to the presence of water and free fatty acids. The presence of even small amounts of water can lead to saponification, an undesirable side reaction. Saponification produces soap, which increases the viscosity of the fuel and results in product losses. The formation of soap gels complicates the recovery of the by-product glycerol as well as that of the fatty acid methyl esters. It also consumes the basic catalyst, reducing the rate and yield of the reaction. In oils with high moisture content and free fatty acids, an acid catalyst is often employed. The use of basic catalysts is also energy intensive and can produce a substantial amount of high-pH waste products. Alternatively, triglycerides in the oil can be first saponified. The free fatty acid solution can then be purified (H_2O removed) and then transesterified with an acid catalyst.

Numerous studies have been published on the kinetics of the transesterification reaction. These studied various parameters that affect the transesterification reaction:

(1) molar ratio of alcohol to vegetable oil (“concentration”), (2) type of vegetable oil (rapeseed oil, soybean oil, waste frying oil), (3) catalyst type (acid, base, enzyme or none), (4) temperature, (5) impurities (presence of free fatty acids, moisture), (6) mixing intensity.

6.4 A Kinetic Model of Chemical Transesterification

Complete transesterification of triglycerides to yield fatty acid methyl esters consists of three sequential, step-wise and reversible reactions. Triglycerides are first converted to diglycerides, which are then converted to monoglycerides, which are eventually converted to glycerol and FAMEs. A FAME is produced at each step. Completion of the reaction consumes a triglyceride and three molecules of alcohol to produce glycerol and three molecules of FAME. The equilibrium of the reaction tends towards the production of FAMEs and glycerol.

The experimental model assumes that the oil contains only triglycerides and that all triglyceride molecule species react at the same rate. The model also assumes that there is no production of soap during the course of the reaction.

Note that TG = Triglyceride, DG = Diglyceride, MG = Monoglyceride, E = Fatty acid ester, A = Alcohol and GL = Glycerol. Transesterification consists of the following elementary reactions:

$$\mathrm{TG} + \mathrm{A} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{DG} + \mathrm{E}$$

$$\mathrm{DG} + \mathrm{A} \underset{k_4}{\overset{k_3}{\rightleftharpoons}} \mathrm{MG} + \mathrm{E}$$

$$\mathrm{MG} + \mathrm{A} \underset{k_6}{\overset{k_5}{\rightleftharpoons}} \mathrm{GL} + \mathrm{E}$$

The overall stoichiometry of the reaction can be written as:

$$\mathrm{TG} + 3\ \mathrm{CH_3OH} \rightleftharpoons \mathrm{GL} + 3\mathrm{E}$$

The following differential equations can be written to describe the rate of the reaction:

$$\frac{d[\mathrm{TG}]}{dt} = -k_1[\mathrm{TG}][\mathrm{A}] + k_2[\mathrm{DG}][\mathrm{E}] \tag{6.1}$$

$$\frac{d[\mathrm{DG}]}{dt} = k_1[\mathrm{TG}][\mathrm{A}] - k_2[\mathrm{DG}][\mathrm{E}] - k_3[\mathrm{DG}][\mathrm{A}] + k_4[\mathrm{MG}][\mathrm{E}] \tag{6.2}$$

$$\frac{d[\mathrm{MG}]}{dt} = k_3[\mathrm{DG}][\mathrm{A}] - k_4[\mathrm{MG}][\mathrm{E}] - k_5[\mathrm{MG}][\mathrm{A}] + k_6[\mathrm{GL}][\mathrm{E}] \tag{6.3}$$

$$\begin{aligned}\frac{d[\mathrm{E}]}{dt} = {} & k_1[\mathrm{TG}][\mathrm{A}] - k_2[\mathrm{DG}][\mathrm{E}] + k_3[\mathrm{DG}][\mathrm{A}] - k_4[\mathrm{MG}][\mathrm{E}] \\ & + k_5[\mathrm{MG}][\mathrm{A}] - k_6[\mathrm{GL}][\mathrm{E}]\end{aligned} \tag{6.4}$$

$$\frac{d[\mathrm{A}]}{dt} = -\frac{d[\mathrm{E}]}{dt} \tag{6.5}$$

$$\frac{d[\mathrm{GL}]}{dt} = k_5[\mathrm{MG}][\mathrm{A}] - k_6[\mathrm{GL}][\mathrm{E}] \tag{6.6}$$

where [A] and [E] denote alcohol and fatty acid methyl ester concentrations. The data was simultaneously fitted to the numerical approximations of the given differential equations. The criterion for goodness-of-fit was the minimized sum of squares. It is important to note here that certain fitting programs allow you to work directly from the differential equations and not the analytical solutions, which can be very convenient. Moreover, some programs also allow for multiple nonlinear regression, in which all equations are fitted simultaneously, obtaining thus a global set of parameter estimates based on the fits of the equations to multiple data sets. The parameter estimates obtained this way are more accurate and robust, but you need greater computational power and more sophisticated programs.

6.4.1 Relating Kinetic Data to Reaction Mechanism

Figure 6.2 shows the concentration of methyl esters over the course of the reaction at 50 °C at different mixing intensities. Alcohol is normally immiscible with triglycerides, resulting in the presence of two phases, an alcohol phase and a triglyceride phase. This implies that the reaction is controlled by the diffusion of reactants in one phase to the other. This limited the rate of the reaction as can be seen by the long lag time observed in the curve describing transesterification at a mixing intensity of 150 rpm. The reaction accelerates, however, once the lag time is past. A probable explanation for this is that FAMEs served to "emulsify" alcohol with the vegetable oil.

Increasing temperature decreases the lag time observed in the mass transfer limited phase. Figure 6.3 shows the effect of temperature on the reaction at 150 rpm. At 150 rpm, the difference in lag time between the reaction at 30 and 70 °C is approximately 35 min.

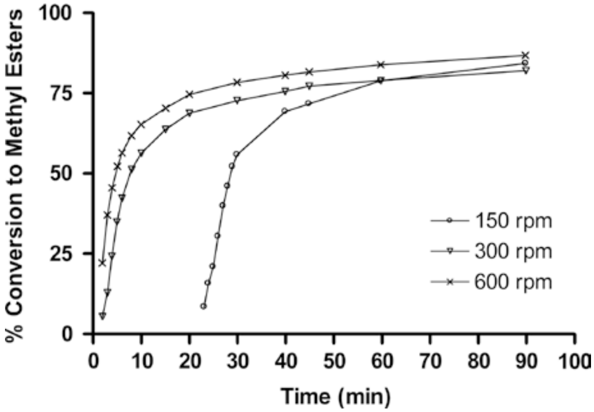


Fig. 6.2 The production of fatty acid methyl esters at 50 °C over time as influenced by three mixing intensities, 150 rpm, 300 rpm and 600 rpm

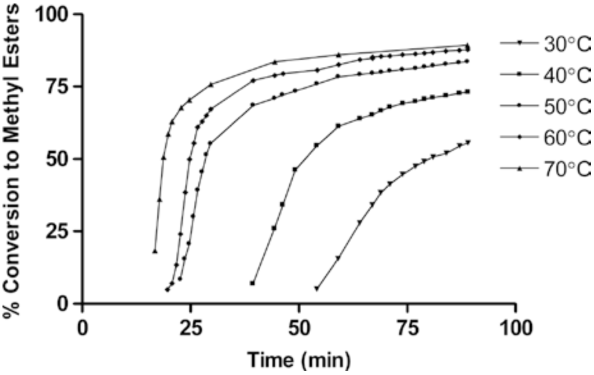


Fig. 6.3 The effect of different temperatures on the transesterification reaction conducted with a mixing intensity of 150 rpm

The effect of agitation speed on the rate of the reaction can be easily visualized by comparing Figs. 6.3 and 6.4, which displays the course of different tranesterification reactions at 300 rpm. The data shows the effect of mixing on the reaction rate. At 30 °C and 150 rpm, the lag time is approximately 55 min. In contrast, at 30 °C and 300 rpm, the lag time is less than 10 min. This trend is also evident at other reaction temperatures, with the mixing rate having less of an effect at higher temperatures. For example, at 150 rpm, the lag time for the reaction at 70 °C is approximately 20 min. The lag time at 300 rpm is approximately 2 min. The difference between the delay time at 70 °C is roughly 18 min. Contrast this to the difference of nearly 45 min for the reaction at 30 °C.

Table 6.1 gives the rate constants alongside the activation energy of each reaction step and its corresponding reverse reaction. It can be seen that a reaction with a higher activation energy has a lower rate constant, indicating a slower reaction. This is in agreement with the Arrhenius relationship. The data indicates that for the conversion of triglycerides to diglycerides, the magnitude of the forward reaction rate constant is less than that of the reverse reaction's. Similarly, the conversion of diglyceride to monoglyceride is less kinetically favorable than the reverse reaction. However, the data also shows that the last reaction step, the conversion of monoglyceride to glycerol is kinetically favorable as evidenced by the lower activation energy as well as a higher reaction rate constant.

Fig. 6.4 The progress of the transesterification reaction at different temperatures and a mixing intensity of 300 rpm

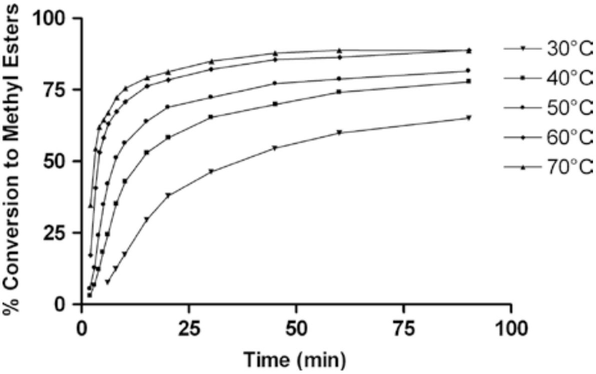


Table 6.1. Energy of activation (kJ/mol) and rate constants (1/(M alcohol)·s) of different reactions at 300 rpm

Reaction	Activation energy kJ/mol	Rate constant (min^{-1})
TG → DG	55.00	0.05
DG → TG	41.56	0.110
DG → MG	83.09	0.215
MG → DG	61.25	1.228
MG → GL	26.87	0.242
GL → MG	40.12	0.007

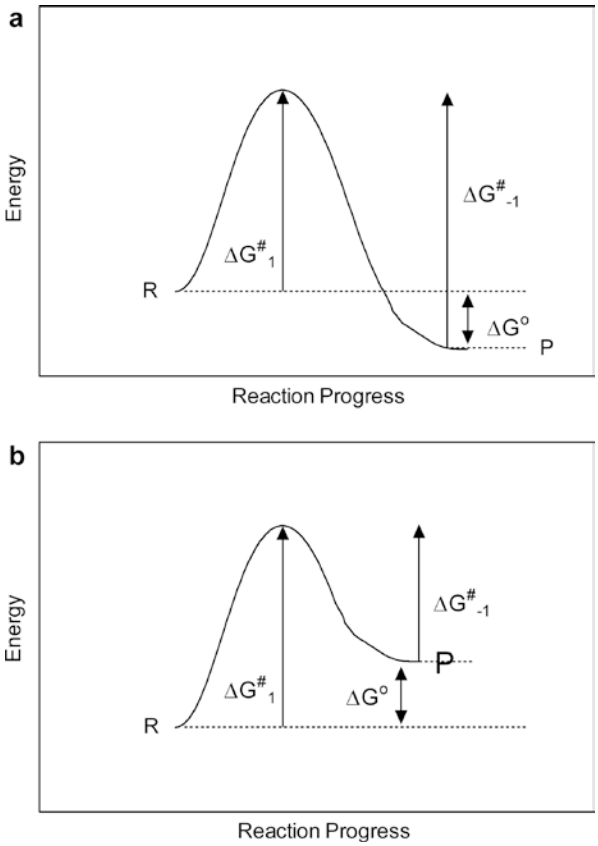


Fig. 6.5 The free energy – reaction progress plots of an exergonic (**a**) and endergonic (**b**) reaction

If each step is assumed to be an elementary reaction step, thermodynamic information on the free energy (ΔG) released or required can be calculated from the activation energies. The free energy change of a reaction in solution can be approximated by substracting the activation energy of the reverse reaction from the activation energy of the forward reaction. A reaction with an activation energy that is higher than the activation energy of the reverse reaction will give a positive ΔG, indicating an exergonic reaction (Fig. 6.5a). Similarly, a reaction with a lower activation energy than that of the reverse reaction is exergonic (Fig. 6.5b). Examining the data, it can be seen that the conversion of triglyceride to diglyceride and of diglyceride to monoglyceride are endergonic reactions. The free energy change, ΔG, of the TG → DG reaction is 13.44 kJ·mol^{-1}. Similarly the free energy change of the DG → MG reaction is 21.84 kJ·mol^{-1}. The conversion of monoglyceride to glycerol is an exergonic reaction as can be surmised from its negative free energy change of -13.25 kJ·mol^{-1}. These calculations indicate that the last step of the reaction is essentially irreversible whereas the first two are reversible.

This example has shown how kinetic modelling can be a very useful tool in the study of processes, and how detailed consideration of the value of the reaction parameters can be used to propose mechanisms for the reaction and/or to better understand it.

Bibliography

1. Noureddini H, Zhu D (1997) Kinetics of transesterification of soybean oil. J Am Oil Chem Soc 74:1457–1463
2. Ma F, Hanna MA (1999) Biodiesel production: a review. Bioresour Technol 70:1–15
3. Srivastava A, Prasad R (2000) Triglycerides-based diesel fuels. Renew Sustain Energy Rev 4:111–133
4. Knothe G, Gerpen JV, Krahl J (2005) The biodiesel handbook. AOCS Press, Urbana

Chapter 7
Kinetics of Crystal Growth Using the Avrami Model and the Chemical Potential Approach

7.1 Introduction

The Avrami model of crystallization was proposed by the scientist Melvin Avrami in the early 1940s. It was originally used to quantify the liquid-solid phase transition kinetics in metals. Its principles have been since expanded to model other systems including fat and starch crystallization. The utility of the model stems from its ability to provide indications on the nature of the crystal growth process. Applied to the study of fat crystallization, the Avrami equation has the following form,

$$\frac{\mathrm{SFC}}{\mathrm{SFC}_{\max}} = 1 - e^{-kt^n} \tag{7.1}$$

where n is the Avrami exponent (dimensionless), *k* is the Avrami constant, SFC corresponds to the solid fat content at time *t*, and $\mathrm{SFC}_{\max}$ corresponds to the maximum SFC achieved at a given temperature.

The Avrami model is the most commonly used model in the study of fat crystallization. The equation describes an event in which there is an initial lag-period, where crystallization occurs very slowly, followed by a subsequent rapid increase in crystal mass. This model takes into account that crystallization occurs by both nucleation and crystal growth. The model makes several assumptions such as isothermal conditions, random nucleation in space, and linear growth kinetics wherein the growth rate of the new phase depends only on temperature and not on time. The density of growing bodies is also assumed to remain constant.

The Avrami parameters provide information on the nature of the crystallization process. The constant, *k*, is the crystallization rate constant. It is mainly a function of the crystallization temperature and generally obeys an Arrhenius-type temperature

A.G. Marangoni, *Kinetic Analysis of Food Systems*,
DOI 10.1007/978-3-319-51292-1_7

Table 7.1 Values for the Avrami exponent, *n*, for different types of nucleation and growth

N	Type of crystal growth and nucleation expected
3 + 1 = 4	3D growth from sporadic nuclei (spheroids)
3 + 0 = 3	3D growth from instantaneous nuclei (spheroids)
2 + 1 = 3	2D growth from sporadic nuclei (discs, plates)
2 + 0 = 2	2D growth from instantaneous nuclei (discs, plates)
1 + 1 = 2	1D growth from sporadic nuclei (needles, rods)
1 + 0 = 1	1D growth from instantaneous nuclei (needles, rods)

Adapted from Sharples (1966)

dependency. The crystallization rate constant takes both nucleation rates and crystal growth rates into account. Half-times of crystallization, $t_{\frac{1}{2}}$, reflect the magnitudes of *k* and *n* according to:

$$t_{\frac{1}{2}} = \left(\frac{\ln 2}{k}\right)^{\frac{1}{n}} \tag{7.2}$$

The Avrami exponent, *n*, sometimes referred to as the index of crystallization, indicates the crystal growth mechanism. This parameter is a combined function of the time dependence of nucleation and the number of dimensions in which growth takes place. Nucleation is either instantaneous, with nuclei appearing all at once early on in the process, or sporadic, with the number of nuclei increasing linearly with time. Growth either occurs as rods, discs, or spheres in one, two, or three dimensions, respectively. Table 7.1 shows the value of the Avrami exponent, *n*, expected for various types of nucleation and growth.

Although *n* should be an integer, fractional values are usually obtained, even in cases where the model fits the data quite well. Deviations from integer values for *n* have been explained as simultaneous development of two (or more) types of crystals, or similar crystals from different types of nuclei (sporadic vs. instantaneous). Deviations may also occur in cases where spherical crystals arise from initially rod- or plate-like nuclei, such as is the case for spherulites.

Following the philosophy of Christian, specific Avrami exponents can be associated with certain growth modes determined by microscopy. For example, milk fat crystallized at a high degree of supercooling will have a lower activation energy for the nucleation process. Since the kinetic barrier is lower, a high rate of nucleation will be observed. This, in turn, will lead to the formation of a large number of nuclei at the onset of the crystallization process (instantaneous nucleation). The rate of crystal growth would also be quite high. This in turn would lead to a more one-dimensional growth. The end result would be a granular microstructure composed of a large number of small crystals as can be seen in Fig. 7.1.

On the other hand, fats crystallized at low degrees of undercooling will have a higher free energy of nucleation and a lower rate of nucleation. A smaller number of

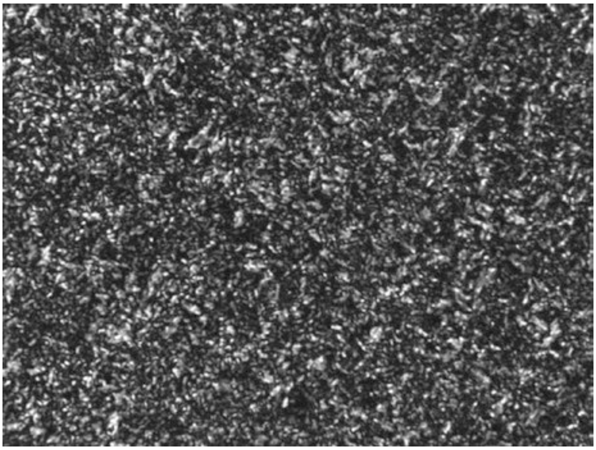


Fig. 7.1 Microstructure of cocoa butter after being crystallized isothermally at 5 °C for 7 days. The Avrami index for the crystallization process leading to this microstructure was $n = 1$

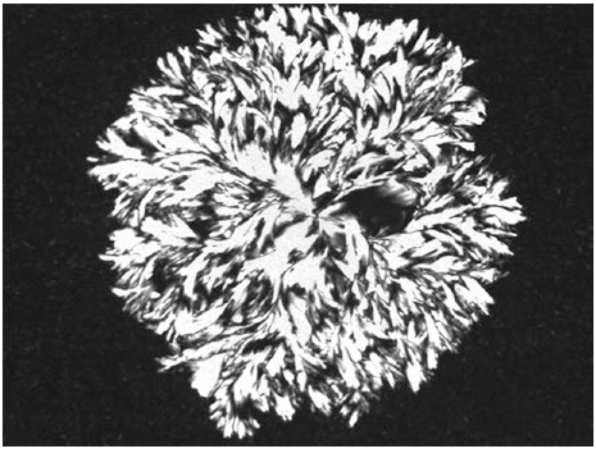


Fig. 7.2 Microstructure of cocoa butter after being crystallized isothermally at 20 °C for 21 days. The Avrami index for the crystallization process leading to this microstructure was $n = 4$

nuclei will be formed, possibly in a sporadic fashion (sporadic nucleation). The rate of crystal growth would also be lower. This in turn would lead to a growth mode that is more multi-dimensional. The end result would be a "clustered" microstructure composed of a small number of large crystals as shown in Fig. 7.2.

7.2 Derivation of the Avrami Model

The starting point for the derivation of the Avrami model is the empirical chemical diffusion equation also called Fick's first law in one dimension.

$$\frac{\partial Q}{\partial t} = KA\left(\frac{\Delta C}{\Delta x}\right) \tag{7.3}$$

Fick's first law of diffusion describes the transport of mass per unit area, $\frac{\partial Q}{\partial t}$, across a concentration difference (mol·m^3), ΔC. Since the rate of increase in mass of a crystal $\left(\frac{\partial m_c}{\partial t}\right)$ is a function of the amount of mass that diffuses to the crystal surface, we can rewrite Eq. 7.3 as the following:

$$\frac{\partial m_c}{\partial t} = \frac{\partial Q}{\partial t} MW \tag{7.4}$$

The growth of a single crystal is proportional to the surface area A_g (in m^2) of the crystal nucleus and the concentration difference, $c - c^*$ (in kg·m^{-3}), between the supersaturated material and the equilibrium concentration of the same material at a specific temperature:

$$\frac{\partial m_c}{\partial t} \propto A_g \left(c - c^*\right) \tag{7.5}$$

A rate constant for single crystal growth k_g (in m·s^{-1}) can be written for this proportionality to obtain the following:

$$\frac{\partial m_c}{\partial t} = k_g A_g \left(c - c^*\right) \tag{7.6}$$

The crystallization rate constant k_g is a function of the mass transfer/diffusion coefficient K (in m^2·s^{-1}), the surface reaction rate constant κ (in m^2·s^{-1}) and the distance/thickness Δx (in m) through which diffusion takes place. The Δx term is also called the boundary layer. The dependence of the crystallization rate constant k_g on these parameters is given in Eq. 7.7.

$$k_g = \left(\frac{1}{K} + \frac{1}{\kappa}\right)^{-1} \frac{1}{\Delta x} \tag{7.7}$$

The concentration difference $c - c^*$ (in terms of kg·m^{-3}) can be rewritten as the concentration gradient in moles per m^{-3} (ΔC) multiplied by the molecular weight MW (in kg·mol^{-1}):

$$\left(\Delta C\right) MW = \left(c - c^*\right) \tag{7.8}$$

The rate constant for single crystal growth (k_g) accounts for the possibility of diffusion control at small relative velocities and surface reaction control at high relative velocities. The diffusion coefficient (K) varies with solution properties like viscosity and agitation, but the surface reaction rate constant (κ) does not. As well, K varies little with temperature, while κ can change dramatically on cooling.

The rate of total increase in solids in the system $\left(\frac{\partial m_s}{\partial t}\right)$ is the product of the increase in mass of a single crystal, $k_g A_g (c - c^*)$, times the number of growing crystals in the system, N_c:

$$\frac{\partial m_s}{\partial t} = N_c k_g A_g \left(c - c^*\right) \tag{7.9}$$

The change in concentration, $c - c^*$ (in kg·m^3), is simply the change in mass divided by the total volume:

$$\left(c - c^*\right) = \frac{\left(m_{\max} - m_s\right)}{V_T} \tag{7.10}$$

By substituting Eq. 7.10 into Eq. 7.9, the following is obtained:

$$\frac{\partial m_s}{\partial t} = N_c k_g A_g \frac{\left(m_{\max} - m_s\right)}{V_T} \tag{7.11}$$

The crystallite density ρ_c (crystals per m^{-3}) is expressed as the number of crystals (N_c) per unit volume (V_T) and is defined as:

$$\rho_c = \frac{N_c}{V_T} \tag{7.12}$$

which can be re-arranged to give $\rho_c V_T = N_c$.

After substitution of Eq. 7.12 into Eq. 7.11 and re-arrangement, the following expression is obtained:

$$\frac{\partial m_s}{\left(m_{\max} - m_s\right)} = k_g \rho_c A_g \partial t \tag{7.13}$$

This is the basic equation that can be used to derive the final form of the Avrami equation for different growth geometries and types of nucleation (i.e., different Avrami exponents).

7.3 Spherical Growth with Instantaneous Nucleation

The surface area A_g (m^2) of a spherical crystal involved in crystal growth is

$$A_g = 4\pi r^2 \tag{7.14}$$

where r is the crystal radius (m). The linear growth rate of the crystal radius in time is expressed as:

$$r = gt \tag{7.15}$$

Introducing Eq. 7.14 and Eq. 7.15 into Eq. 7.13 results in:

$$\frac{\partial m_s}{\left(m_{\max} - m_s\right)} = k_g \rho_c 4\pi g^2 t^2 \partial t \tag{7.16}$$

Integration of the above expression for the boundary conditions $m_s = 0$ at $t = 0$, and m_s at t,

$$\int_0^{m_s} \frac{\partial m_s}{\left(m_{\max} - m_s\right)} = 4\pi k_g \rho_c g^2 \int_0^t t^2 \partial t \tag{7.17}$$

results in the equation

$$\ln\left(\frac{m_{\max}}{m_{\max} - m_s}\right) = \frac{4}{3}\pi k_g \rho_c g^2 t^3 \tag{7.18}$$

This expression can be transformed and rearranged to:

$$\frac{m_s}{m_{\max}} = 1 - e^{-\frac{4}{3}\pi k_g \rho_c g^2 t^3} \tag{7.19}$$

Moreover, the mass fraction $\left(\frac{m_s}{m_{\max}}\right)$ can be expressed as the ratio of the solid fat content at a particular time SFC to the maximum solid fat content as $t \to \infty$, $SFC_{\max}$:

$$\frac{m_s}{m_{\max}} = \frac{\frac{m_s}{m_T}}{\frac{m_{\max}}{m_T}} = \frac{SFC}{SFC_{\max}} \tag{7.20}$$

Thus, the Avrami equation describing the growth of a spherical crystal under conditions of instantaneous nucleation can be expressed as:

$$\frac{SFC}{SFC_{\max}} = 1 - e^{-k_A t^3} \tag{7.21}$$

where

$$k_A = \frac{4}{3}\pi k_g \rho_c g^2 \tag{7.22}$$

7.3.1 Spherical Growth with Sporadic Nucleation

The treatment starts with (Eq. 7.16), shown here again for the sake of clarity,

$$\frac{\partial m_s}{(m_{\max} - m_s)} = k_g \rho_c 4\pi g^2 t^2 \partial t \tag{7.16}$$

For sporadic nucleation, the change in the number of nuclei as a function of time is given as:

$$\rho_c = \frac{N_c}{V_T} = jt \tag{7.23}$$

Introducing Eq. 7.23 into Eq. 7.16 results in:

$$\frac{\partial m_s}{(m_{\max} - m_s)} = k_g j 4\pi g^2 t^3 \partial t \tag{7.24}$$

Integration of the above expression for the boundary conditions $m_s = 0$ at $t = 0$, and m_s at t,

$$\int_0^{m_s} \frac{\partial m_s}{(m_{\max} - m_s)} = 4\pi k_g j g^2 \int_0^t t^3 \partial t \tag{7.25}$$

results in the expression:

$$\ln\left(\frac{m_{\max}}{m_{\max} - m_s}\right) = \pi k_g j g^2 t^4 \tag{7.26}$$

which can be transformed and rearranged to give:

$$\frac{m_s}{m_{\max}} = 1 - e^{-\pi k_g j g^2 t^4} \tag{7.27}$$

Thus, the Avrami equation describing the growth of a spherical crystal under conditions of sporadic nucleation can be expressed as

$$\frac{SFC}{SFC_{\max}} = 1 - e^{-k_A t^4} \tag{7.28}$$

where

$$k_A = \pi k_g j g^2 \tag{7.29}$$

7.3.2 *Plate-Like Growth with Instantaneous Nucleation*

Consider a rectangular plate growing in the X-Y plane, but not in the Z-plane. The area involved in crystal growth A_g is therefore:

$$A_g = 4lh \tag{7.30}$$

where the dimension l is increasing linearly in time according to

$$l = gt \tag{7.31}$$

Introducing Eqs. 7.30 and 7.31 into Eq. 7.16 with subsequent integration and rearrangement results in:

$$\frac{m_s}{m_{\max}} = \frac{SFC}{SFC_{\max}} = 1 - e^{-k_A t^2} \tag{7.32}$$

where

$$k_A = 2hk_g \rho_c g \tag{7.33}$$

7.4 Plate-Like Growth with Sporadic Nucleation

For sporadic nucleation, the change in the number of nuclei as a function of time is given in Eq. 7.23, shown here again for the sake of clarity,

$$\rho_c = \frac{N_c}{V_T} = jt \tag{7.23}$$

Introducing Eq. 7.23 into Eq. 7.16 along with integration and rearrangement results in:

$$\frac{m_s}{m_{\max}} = 1 - e^{-\frac{4}{3}hk_g jgt^3} \tag{7.33}$$

Thus, the Avrami equation describing the growth of a plate-like crystal in 2-dimensions under conditions of sporadic nucleation can be expressed as

$$\frac{m_s}{m_{\max}} = \frac{SFC}{SFC_{\max}} = 1 - e^{-k_A t^3} \tag{7.34}$$

where

$$k_A = \frac{4}{3} h k_g jg \tag{7.35}$$

7.4.1 Rod-Like Growth with Instantaneous Nucleation

Consider a cylinder growing in length, but not in cross-section. The area involved in crystal growth is the area of the base ends:

$$A_g = 2\pi r^2 \tag{7.36}$$

Introducing Eq. 7.36 into Eq. 7.16 along with integration and rearrangement results in:

$$\frac{m_s}{m_{\max}} = \frac{SFC}{SFC_{\max}} = 1 - e^{-k_A t} \tag{7.37}$$

where

$$k_A = 2\pi r^2 k_g \rho_c \tag{7.38}$$

7.4.2 Rod-Like Growth with Sporadic Nucleation

For sporadic nucleation, the change in the number of nuclei as a function of time is given by Eq. 7.23, shown here again for the sake of clarity,

$$\rho_c = \frac{N_c}{V_T} = jt \tag{7.23}$$

Introducing Eqns. 7.23 and 7.36 into Eq. 7.16, integration and rearrangement results in

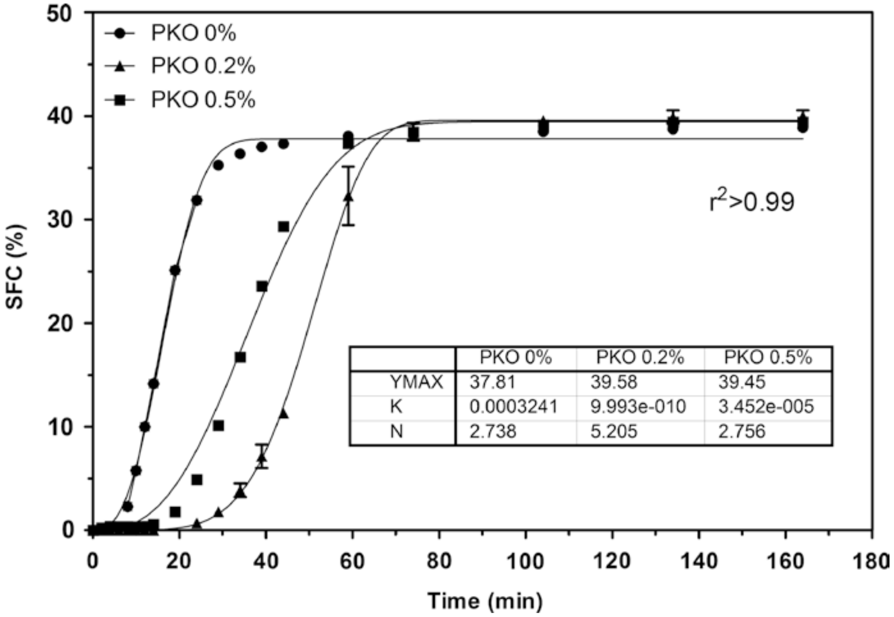
$$\frac{m_s}{m_{\max}} = \frac{SFC}{SFC_{\max}} = 1 - e^{-k_A t^2} \tag{7.39}$$

where

$$k_A = \pi r^2 k_g j \tag{7.40}$$

7.5 Transformation and Use of the Model in Experiments

SFC vs. time data can be fitted to the Avrami model in its nonlinear form (Eq. 7.1), using standard curve-fitting (nonlinear regression) routines available in most modern data analysis programs, with estimates of k_A and n being obtained. Model fits to experimental data and estimates of the kinetic parameters are shown in Fig. 7.3. This data set corresponds to the isothermal crystallization of palm kernel oil (PKO) in the presence or absence of different amounts of a "stabilizer", a blend of fully hydrogenated high-erucic acid rapeseed oil and soybean oil. According to this analysis, PKO crystallization is strongly affected by intermediate levels of stabilizer addition (0.2%), which are reversed upon addition of higher amounts (0.5%). The Avrami index suggests possibly a plate-like growth with sporadic nucleation or a spherical growth with instantaneous nucleation for PKO without stabilizer and 0.5% stabilizer. Addition of 0.2% stabilizer slows down the reaction considerably and increases the Avrami index to about n = 5. There is no interpretation for n > 4 in the Avrami model, and one must be very careful with this since the Avrami index can be affected by non-crystallization-related factors such as heat and mass transfer limitations and crystallizer geometry, resulting in a long lag time prior to commencement of crystallization.



	PKO 0%	PKO 0.2%	PKO 0.5%
YMAX	37.81	39.58	39.45
K	0.0003241	9.993e-010	3.452e-005
N	2.738	5.205	2.756

Fig. 7.3 Change in solid fat content as a function of time for the isothermal crystallization of palm kernel oil (*PKO*) in the presence or absence of a stabilizer composed of fully hydrogenated high erucic acid rapeseed oil and soybean oil. Data fits to the Avrami model and parameter estimates are shown

As stated above, the Avrami model does not implicitly take induction times into consideration. However, if induction times (t_o) are somehow independently determined, these can be incorporated into the Avrami model:

$$\frac{SFC}{SFC_{\max}} = 1 - e^{-k_A(t-t_o)^n} \tag{7.41}$$

This is particularly important point since a process with an actual induction time, as evidenced in the Fig. 7.3 for PKO 0.2%, cannot be properly modelled using the Avrami model. This delay could be, for example, due to slow cooling effects and poor mixing. In the experience of the author, Avrami exponents higher than four are most probably artefactual and the model should then include such induction time. However, such induction time should be determined independently since estimating it as a fitting parameter upon curve fitting will most probably introduce large errors in the overall fit of the model to the data. The induction time needs to be fixed as a constant.

If nonlinear regression is not available to the researcher, it is sometimes possible to linearize the model and use linear regression instead using a variety of more widely available programs. We outline this procedure below and provide an example.

The Avrami model has the general form

$$\frac{SFC}{SFC_{\max}} = 1 - e^{-k_A t^n} \tag{7.1}$$

This model can be linearized by a double logarithmic transformation to give:

$$\ln\left(1 - \frac{SFC}{SFC_{\max}}\right) = -k_A t^n \tag{7.42}$$

$$\ln\left(-\ln\left(1 - \frac{SFC}{SFC_{\max}}\right)\right) = \ln\left(k_A\right) + n\ln\left(t\right) \tag{7.43}$$

Thus, a plot of $\ln\left(-\ln\left(1 - \frac{SFC}{SFC_{\max}}\right)\right)$ vs. ln (t) should yield a straight line with a slope n and a y-intercept of $\ln(k_A)$. The same PKO data shown in Fig. 7.3 is shown in its "linearized" form in Fig. 7.4. We would expect the linearization to yield a straight line when plotted. However, the transformed points do not yield a straight line! Interestingly, for PKO with 0.2 and 0.5% stabilizer added, linear regions can be discerned. The situation for PKO with no stabilizer (0%) is more complicated. Notice that the estimates for the Avrami index and Avrami constant are not equivalent to those obtained by nonlinear regression. A lack of a linear region may be indicative of a variable index and constant, or some other unforeseen factor. It is the experience of the author that regardless of the fact that linearization of crystallization data is widely used, it is best to stick to nonlinear regression, which uses the entire data set. Another

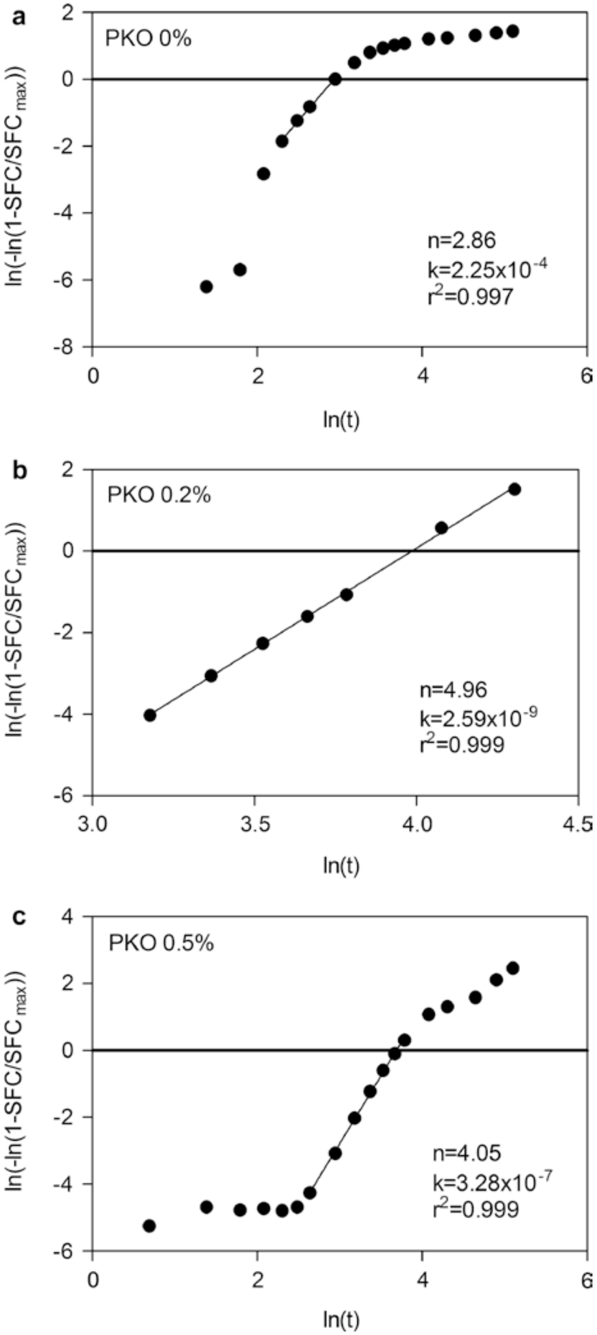


Fig. 7.4 Linearization of the SFC-temperature data presented in Fig. 7.3 for each treatment in each separate panel for determination of the Avrami kinetic parameters by linear regression a linear region within the transformed data. Estimates of the kinetic parameters form the linear regression fits indicated in the data are shown

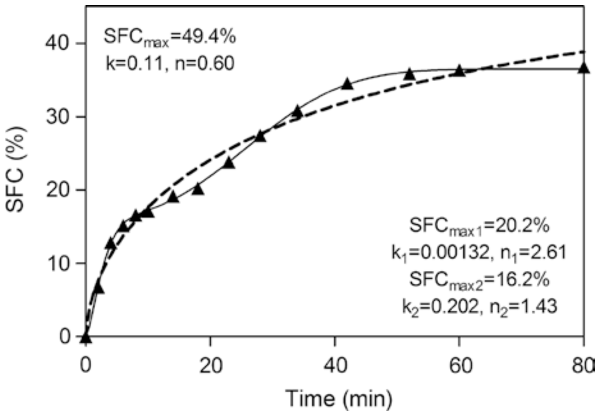


Fig. 7.5 Evolution of the solid fat content (*SFC*) in time during the isothermal crystallization of milkfat at 10 °C. Shown here are single-step (*dashed line*) and multiple step (*solid line*) Avrami model fits to the data. Estimates of the kinetic parameters for each fit are indicated as well

limitation of linearization is that SFC$_{\max}$ must be chosen initially in order to normalize the data, and many times this SFC$_{\max}$ does not correspond to the limiting value at time approaches infinity. Most the time, the last value in the data set is taken as the SFC$_{\max}$, which may not be correct.

Another variation on the Avrami models arises because multiple-step growth curves are observed in crystallization experiments. Multiple steps may be due to fractionation or polymorphic transformations taking place during crystallization. For example, the SFC-time profile of milkfat being crystallized isothermally at 10 °C is given in Fig. 7.5. The entire process could be modeled using a multi-component summation of the Avrami model of the form:

$$SFC = SFC_o + \sum_{i=1}^{n} SFC_{\max,i}\left(1 - e^{-k_i t^{n_i}}\right) \tag{7.44}$$

where SFC$_o$ is the initial SFC, usually zero. The parameter estimates are shown in Fig. 7.5 and characterize the two steps involved in the step-wise solidification of milkfat under static conditions at this temperature.

If the multi-component model is not used and the data is instead fitted as a single component Avrami equation (Eq. 7.1), a different set of kinetic parameters is obtained as can be seen from Fig. 7.5. The single-component analysis can be taken to represent the average of crystallization events and is not representative of what is going on.

7.6 An Alternative to the Avrami Model: the Chemical Potential Approach

In this chapter I would like to present an alternative, and probably more proper, way to characterize the kinetics of crystallizations of crystalline materials, fats in particular. The Avrami model is a very powerful tool to characterize the kinetics of crystallization. However, as we will see below, the model is not based on a “driving force”

Table 7.2 Crystallization and melting temperatures (onset and peak) for 40/60% PS-SO, 20/80% FHSO-SO and 20/80% FHSO-HOSFO oil blends determined by differential scanning calorimetry using a rate of 5 °C/min

System	Melting temperature (Onset)	Crystallization temperature (Onset)	Melting temperature (peak)	Crystallization temperature (peak)
40/60 PS-SO	44.8 ± 6.4	19.6 ± 0.3	50.6 ± 0.1	18.4 ± 0.1
20/80 FHSO-SO	54.8± 3.5	39.1 ± 0.2	60.9 ± 0.6	37.1 ± 0.6
20/80 FHSO-HOSFO	54.2± 2.2	37.4 ± 0.1	54.5 ± 1.6	36.4 ± 0.1

Values represent averages and standard deviations of three determinations

for crystallization and the growth curves obtained can be riddled with artifacts related to polymorphic transformations during the course of crystal growth, and heat-transfer limitations during cooling. Moreover, the model is supposed to be used for one-component systems, but is extensively used for multicomponent systems. These artifacts will strongly influence the values obtained for the Avrami index (n) and Avrami constant (k). Another annoying aspect of the Avrami model is that the rate constants are in units of t^{-n}, which makes is very difficult to compare even between treatments of the same sample, and makes the study of the temperature dependence of rate constants for determination of the energy of activation basically impossible.

Here I will analyze the crystallization behavior of three systems, namely 20% fully hydrogenated soybean oil in soybean oil (FHSO-SO), 40% palm stearin in soybean oil (PS-PO) and 20% fully hydrogenated soybean oil in high oleic sunflower oil (FHSO-HOSO). The melting and crystallization temperatures of these systems are given in Table 7.2

Take for example, Figs. 7.6, 7.7 and 7.8. The first two systems seem to be crystallizing in a fashion that would allow their growth curves to be analyzed using the Avrami model. Indeed, as indicated in the figures, fits of the model to the data were excellent and parameters estimated successfully. Even visual inspection of the fitted lines to the experimental points seems outstanding. However, some peculiarities can also be observed. For Fig. 7.6, it would seem that the fully hydrogenated soybean oil in soybean oil accumulates solids initially very rapidly, which are then slowly “melted” away in time, i.e., the SFC-time curve has a maximum. Notice the discrepancies in the maximum SFC predicted by the Avrami model and the experimentally determined one. This was most probably due to the crystallization of a metastable form and slow transformation to a more stable crystal form, along with some melting of the crystals. This was never considered in the Avrami model, and thus we are left with no explanation of why or how this is taking place. To make matters worse, if we had not measured the SFC for a very long time, we would have never been aware of this peculiar behavior. The FHSO-SO samples crystallized in the water bath peak at ~20% solids and decreased all the way to ~16% over 3000 min, while the SFC of samples crystallized in the incubator peaked at ~18% and decreased to ~15% over the 3000 min. The Avrami fits were done for only the first 30 min of crystallization! For palm stearin in soybean oil the situation was better. As shown in Fig. 7.7, the

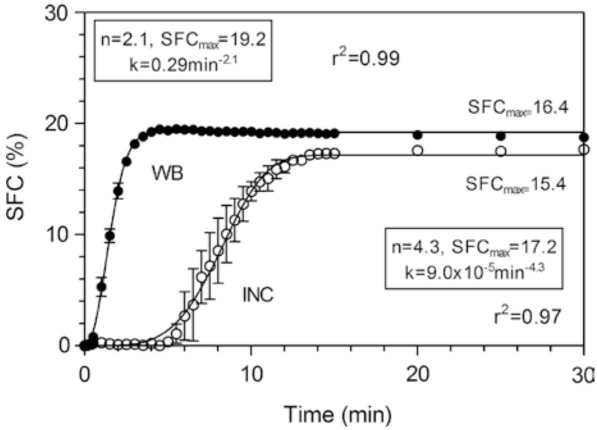


Fig. 7.6 Evolution of the solid fat content (*SFC*) in time during the crystallization of 20% fully hydrogenated soybean oil in soybean oil either in an incubator (*INC*) or in a waterbath (*WB*) set at 30 °C. This translated to approximately ~5 °C/min (*INC*) or ~30 °C/min (*WB*) cooling rates. Indicated at the end of the profiles shown are the maximum equilibrium SFC values determined experimentally. In the boxes we present parameter estimates obtained from the fits of the data to the Avrami model, and next to the box, the r^2 values indicating the goodness of fit

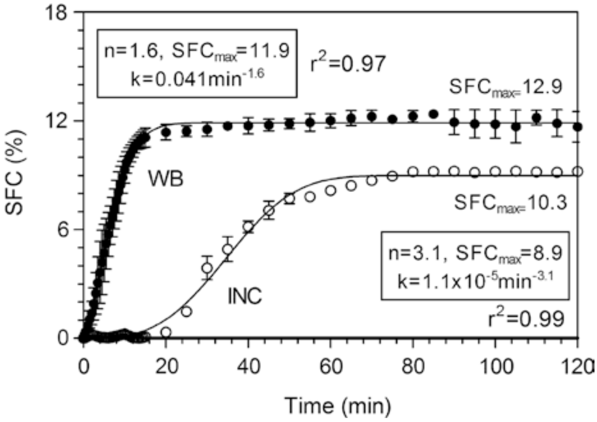


Fig. 7.7 (**a**) Evolution of the solid fat content (SFC) in time during the crystallization of 40% palm stear in soybean oil either in an incubator (*INC*) or in a waterbath (*WB*) set at 20 °C. This translated to approximately ~5 °C/min (*INC*) or ~30 °C/min (*WB*) cooling rates. Indicated at the end of the profiles shown are the maximum equilibrium SFC values determined experimentally. In the boxes we present parameter estimates obtained from the fits of the data to the Avrami model, and next to the box, the r^2 values indicating the goodness of fit

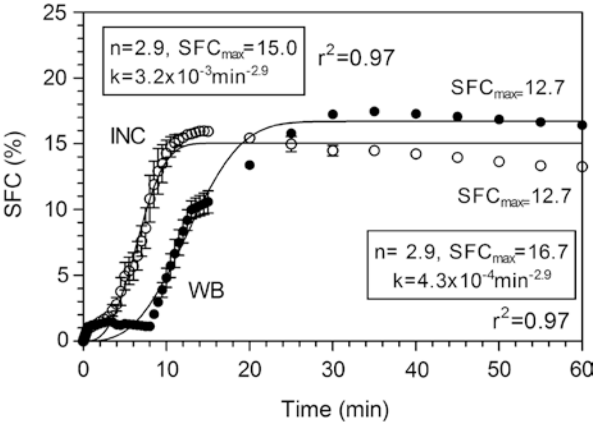


Fig. 7.8 Evolution of the solid fat content (*SFC*) in time during the crystallization of 20% fully hydrogenated soybean oil in high oleic sunflower oil, either in an incubator (*INC*) or in a waterbath (*WB*) set at 40 °C. This translated to approximately ~5 °C/min (*INC*) or ~30 °C/min (*WB*) cooling rates. Indicated at the end of the profiles shown are the maximum equilibrium SFC values determined experimentally. In the boxes we present parameter estimates obtained from the fits of the data to the Avrami model, and next to the box, the r^2 values indicating the goodness of fit

model fit the data well, estimates were reasonable and there was not too much discrepancy between maximum SFC values obtained experimentally and via model fits. For fully hydrogenated soybean oil in high oleic sunflower oil (Fig. 7.8), the situation was very different. This sample differed from FHSO-SO in two aspects, the oil was changed from soybean oil to high oleic sunflower oil (HOSO) and the crystallization temperature was increased from 30 to 40 °C. We determined SO viscosity to be 42 mPa s at 30 °C, while HOSO viscosity was 35 mPa s at 40 °C. I am mentioning this here because the effects noticed when going form SO to HOSO are usually associated with either more viscous crystallization media or lower temperatures, but here temperatures were higher and viscosity was lower. The unusual behavior observed was thus related to the nature of the HOSO oil itself. Figure 7.8 demonstrates further the nature of the problem. Notice how both the samples crystallized in the incubator and in the water bath go through a massive overshoot in solids' accumulation, relative to the final SFC, followed by a marked decrease in time. The fits of the Avrami model to the data are visually poor, but still have a high r^2 value! Notice also the small "step" in the data in the early stages of crystallization. Obviously, the Avrami model cannot account for this crystal transformation and just drives the fit through the average of the data. Nevertheless, some form of quantification of the kinetics of the process was achieved. It depends a bit on what the modeler wants to achieve. However, bear with me and let's analyze this same data from a slightly different perspective.

The driving force for crystallization is the chemical potential difference between the crystallizing molecules in the solid state and in the liquid state. It can be understood best form a supersaturation argument. We have all prepared saturated sugar solutions at some point in our professional career. We can dissolve quite a bit of sugar in water until a specific temperature-dependent concentration, where the solution cannot take in any more solute, and the sugar precipitates out of solution. Let's assume we make such sugar solution at 50 °C and manage to dissolve 260 g sugar in 100 ml of water. If we decrease the temperature to 5 °C, a lot of the sugar that was dissolved at 50 °C will precipitate out, since at 5 °C only 180 g of sugar can be dissolved in 100 ml of water. If we define the saturation concentration as c^*, and the supersaturation concentration of the solution as c, we can define a supersaturation parameter, β, as $\beta = \frac{c}{c^*}$. When we placed our hot sugar solution in the fridge, the supersaturation at that point was 1.44, with a $\ln \beta$ of 0.36. So, we refer to this solution as being "supersaturated", and it definitely needs to lose some of its "energy" in the form of crystallization. This supersaturation is the driving force for crystallization and can be defined in thermodynamic terms as the "chemical potential", $\Delta\mu$, with units of *J/mol*. The relationship between the supersaturation and the chemical potential is given by:

$$\Delta\mu = -RT \ln \beta \tag{7.45}$$

where R is the Universal Gas Constant ($8.314\ J\ mol^{-1}\ K^{-1}$) and T the absolute temperature (*K*).

Now, how do we determine the supersaturation for a crystallizing fat? That is the crux of the problem. In Eq. 7.46, I attempt this by defining the supersaturation of a crystallizing fat as:

$$\beta(T,t) = \frac{c_i(T,t)}{c_i^*(T,eq)} = \frac{\Phi_{neat}^{(T,eq)} - \Phi_{sol}^{(T,t)}}{\Phi_{neat}^{(T,eq)} - \Phi_{sol}^{(T,eq)}} = \frac{x \cdot SFC_{neat}^{(T,\,eq)} - SFC_{sol}^{(T,\,t)}}{x \cdot SFC_{neat}^{(T,\,eq)} - SFC_{sol}^{(T,\,eq)}} \tag{7.46}$$

where *c(T, t)* is the concentration of the supersaturated solution at a specific temperature and/or time, *c*(T, eq)* is the saturation concentration at a specific temperature but when the system has fully reached equilibrium. The "*i*" subscript denotes the *i*th component, should we have more than one component crystallizing. In our analysis, we have considered the fat mixture as a one-component system, which is very convenient and very dangerous. Luckily in fats and oils research, we have a widely-used machine that allows us to determine the solid fat content (SFC) of a fat in about 3 s. So, in this equation, the product of the mass fraction of fat added to an oil (x) and the SFC of the neat fat at equilibrium (remember it usually is not 100% solid) is the maximum amount of solids we would expect to obtain in our fat-oil solution ($x \cdot SFC_{neat}^{(T,\,eq)}$). The difference between this amount and the SFC of the fat-oil solution that we measure at a particular time at a constant temperature ($SFC_{sol}^{(T,\,t)}$), corresponds to the concentration of the supersaturated solution. The difference between $x \cdot SFC_{neat}^{(T,\,eq)}$ and the maximum *SFC* determined for the fat in oil at equilibrium ($SFC_{sol}^{(T,\,eq)}$) corresponds to the saturation concentration. In Eq. 7.46, we also express the equation in more general term using Φ as the volume fraction of solids.

We transformed all the crystallization data in Figs. 7.6, 7.7 and 7.8 using this equation and results are shown in Figs. 7.9a, 7.10a and 7.11a. Notice the complex dynamics observed for FHSO-SO in Fig. 7.9. Since we had an overshoot in the SFC-time

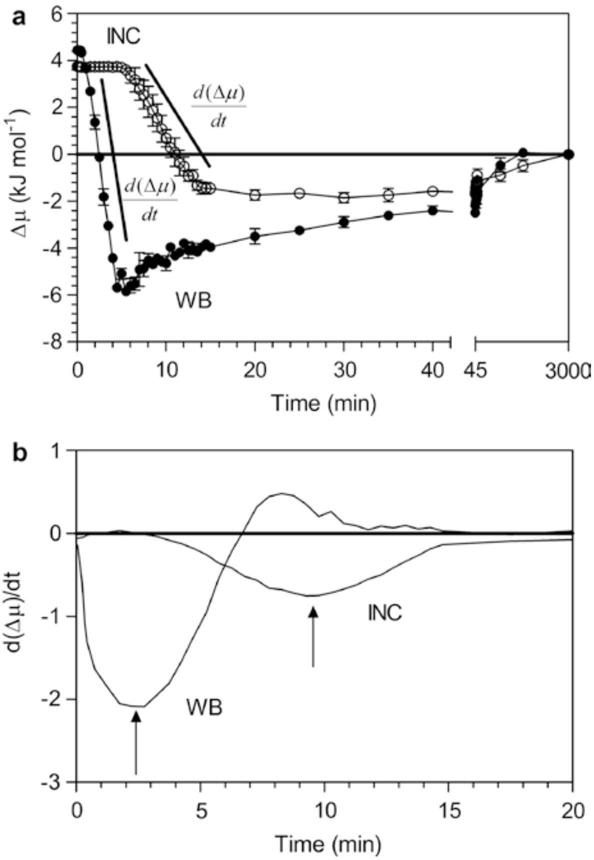


Fig. 7.9 (**a**) Evolution of the chemical potential (Δμ) in time during the crystallization of 20% fully hydrogenated soybean oil in soybean oil either in an incubator (*INC*) or in a waterbath (*WB*) set at 30 °C. This translated to approximately ~5 °C/min (*INC*) or ~30 °C/min (*WB*) cooling rates. Indicated are the general areas used to determine the rate of crystallization from the rate of decrease in chemical potential in time. (**b**) First derivative of the plots in panel *(A)* showing how the maximum rate of chemical potential decrease was determined. We used a 10 neighbor-2nd order smoothing function of the data

patterns (Fig. 7.6.), this translates into a negative chemical potential due to decreases in SFC in time. One can observe here the delay in crystallization for the samples placed in the incubator (air-cooled) vs. the samples placed in a water batch (water-cooled). Also evident is the linear region for loss of chemical potential as a function of time. This is the period of maximum crystallization, and a slope can be easily obtained by linear regression. However, I decided to plot the first derivative of the curve instead (Fig. 7.9b). The minima in the derivative plots correspond to the steepest slopes of the lines. Notice how the system fights hard to get back to a zero chemical potential after a long time. The system will not stop until it reaches this point and the crystallization process is over. Figure 7.10 depicts the case for palm stearin in soybean oil, while Fig. 7.11 depicts the case for fully hydrogenated soybean oil in high oleic sunflower oil.

At this stage, I would like to discuss the reason for my enthusiasm for this linear region. As indicated above and in the figures, this linear region corresponds to $\frac{d(\Delta\mu)}{dt}$. This is the velocity of the crystallization reaction (v). However, let's express this velocity as a function of the supersaturation (β) instead. Going back to Eq. 7.45, the differential of this equation corresponds to:

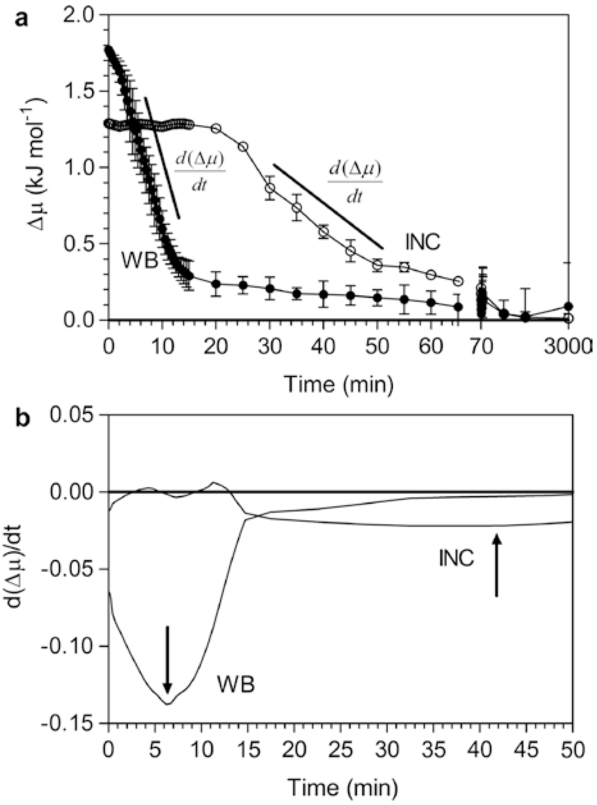


Fig. 7.10 (**a**) Evolution of the solid fat content (*SFC*) in time during the crystallization of 40% palm stear in soybean oil either in an incubator (*INC*) or in a waterbath (*WB*) set at 20 °C. This translated to approximately ~5 °C/min (*INC*) or ~30 °C/min (*WB*) cooling rates. (**b**) First derivative of the plots in panel *(A)* showing how the maximum rate of chemical potential decrease was determined. We used a 10 neighbor-2nd order smoothing function of the data

$$v = \frac{d(\Delta\mu)}{dt} = -RT\frac{d\ln\beta}{dt} \tag{7.47}$$

The rate of the reaction in terms of the rate of change in the supersaturation is simply given by

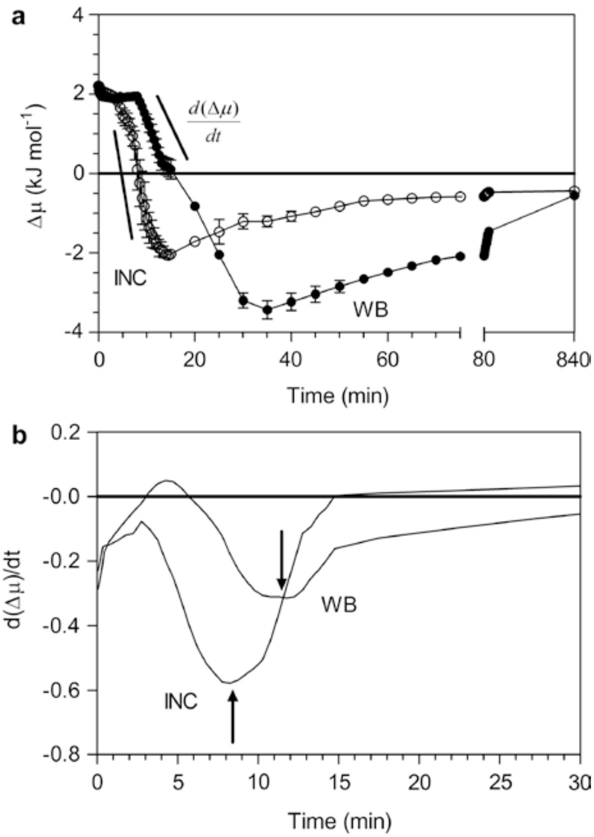
$$\frac{d\ln\beta}{dt} = -\frac{1}{RT}\frac{d(\Delta\mu)}{dt} \quad \left[t^{-1}\right] \tag{7.48}$$

Notice how the units of this rate are [time^{-1}], units of a first order rate constant. This would make comparisons between samples and studies of the temperature dependence of this rate constant much more meaningful.

Moreover, if we integrate Eq. 7.48 between initial time (t_o) and final time (t) for initial supersaturation β, and final supersaturation $\beta = 1$,

$$\int_{\beta}^{1} d\ln\beta = -k_{\beta}\int_{t_0}^{t} dt \tag{7.49}$$

Fig. 7.11 (**a**) Evolution of the chemical potential ($\Delta\mu$) in time during the crystallization of 20% fully hydrogenated soybean oil in high oleic sunflower oil either in an incubator (*INC*) or in a waterbath (*WB*) set at 40 °C. This translated to approximately ~5 °C/min (*INC*) or ~30 °C/min (*WB*) cooling rates. Indicated are the general areas used to determine the rate of crystallization from the rate of decrease in chemical potential in time. (**b**) First derivative of the plots in panel *(A)* showing how the maximum rate of chemical potential decrease was determined. We used a 10 neighbor-2nd order smoothing function of the data



we obtain a simple kinetic expression for the change in supersaturation as a function of time for a crystallizing fat, namely

$$\ln\beta = k_\beta\left(t - t_0\right) \tag{7.50}$$

where $k_\beta = \frac{1}{RT}\frac{d(\Delta\mu)}{dt}$.

It would seem that the time dependence of $ln\beta$ during crystallization is a simple linear function of time. Further studies need to be carried out to test the validity of this new kinetic approach for both isothermal and non-isothermal cases.

This approach was used in the quantification of the kinetic data presented in Figs. 7.9, 7.10 and 7.11. Table 7.3 shows both the estimates of k_β and $\frac{d(\Delta\mu)}{dt}$. As expected, the rates of crystallization were greater for sample placed in a water bath vs. incubator due to cooling rate issues. However, the opposite trend was observed for FHSO-HOSO. This was due to the complexity of the crystallization process

Table 7.3 Rates ($d(\Delta\mu)/dt$) and first order rate constants (k_β) of crystallization derived using the Chemical Potential Approach for 20% Fully Hydrogenated Soybean Oil in soybean oil at 30 °C, 20% Fully Hydrogenated Soybean Oil in High Oleic Sunflower oil at 40 °C and 40% Palm Stearin in Soybean Oil at 20 °C

		Incubator	Water bath
FHSO – SO	$-d(\Delta\mu)/dt$	0.75 $J\ mol^{-1} min^{-1}$	2.1 $J\ mol^{-1} min^{-1}$
	k_β	$3.0 \times 10^{-4}\ min^{-1}$	$8.3 \times 10^{-4}\ min^{-1}$
FHSO – HOSO	$-d(\Delta\mu)/dt$	0.58 $J\ mol^{-1} min^{-1}$	0.31 $J\ mol^{-1} min^{-1}$
	k_β	$2.2 \times 10^{-4}\ min^{-1}$	$1.2 \times 10^{-4}\ min^{-1}$
PS-PO	$-d(\Delta\mu)/dt$	0.022 $J\ mol^{-1} min^{-1}$	0.14 $J\ mol^{-1} min^{-1}$
	k_β	$9.0 \times 10^{-6}\ min^{-1}$	$5.7 \times 10^{-5}\ min^{-1}$

Samples were crystallized either in a waterbath (cooling rate of ~30 °C/min) or incubator (cooling rate of ~5 °C/min)

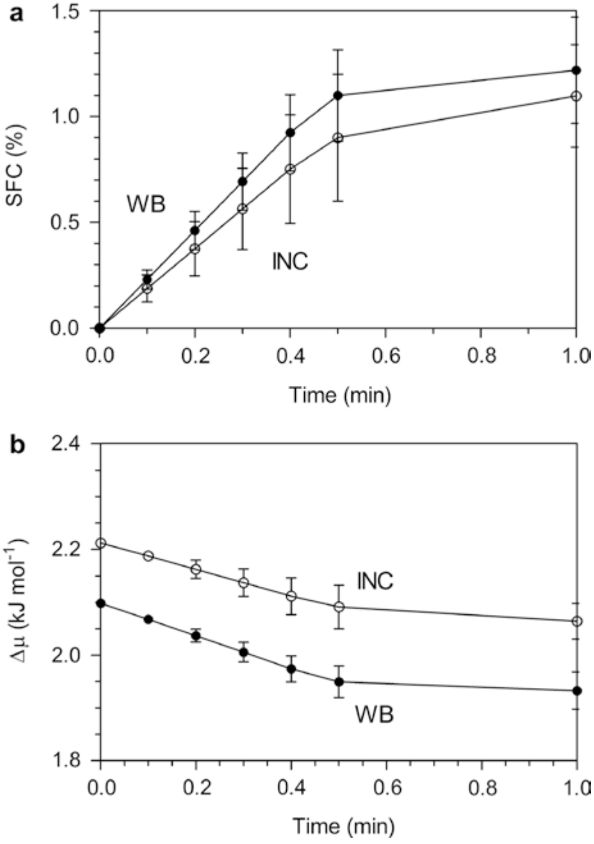


Fig. 7.12 Evolution of the solid fat content (*SFC*) and chemical potential ($\Delta\mu$) in time during the very early crystallization stages of 20% fully hydrogenated soybean oil in high oleic sunflower oil either in an incubator (*INC*) or in a waterbath (*WB*) set at 40 °C. This translated to approximately ~5 °C/min (*INC*) or ~30 °C/min (*WB*) cooling rates

under high supersaturation conditions. Recall the SFC-time profiles showed a "step" in the crystallization process. Upon closer examination of these early stages in the crystallization process (Fig. 7.12), we notice that both SFC-t and $\Delta\mu - t$ profiles are steeper for the waterbath-cooled samples, as expected. What probably happened is that the waterbath-cooled samples crystallized initially very rapidly and extensively into a metastable state, and then upon transformation to a more stable state, released greater amounts of heat than the incubator-cooled samples, and thus slowing down the second step in the crystallization process. In reality we have a transient state followed by a less transient state, followed by a, hopefully, final stable state. Maybe we are cycling through all possible polymorphs in the fat, alpha, beta prime and beta, during the crystallization process. As a matter of fact, we determined FHSO in both HOSO and SO to end up in a beta polymorphic state, but that does not address at all the dynamics of getting to this final beta polymorphic state. One would need time-resolved X-ray diffraction to do this. However, this discussion has highlighted that we need a more complete model to describe the kinetics of polycrystalline materials that undergo metastable nucleation – another opportunity for kinetic modelers to jump in and have fun!

This chapter has helped introduce the student of crystallization to the well-known formalism of the Avrami model, while at the same time demonstrating the power of kinetic modelling in helping to better understand crystallization using a Chemical Potential Approach. As stated in the introduction to this book, this new approach/model is a hypothesis that we can entertain.

Bibliography

1. Avrami M (1939) Kinetics of phase change I general theory. J Chem Phys 7:1103–1112
2. Avrami M (1940) Kinetics of phase change II. Transformation-time relations for random distribution of nuclei. J Chem Phys 8:212–224
3. Avrami M (1941) Kinetics of phase change III. Granulation, phase change, and microstructure. J Chem Phys 9:177–184
4. Christian JW (1965) The theory of transformations in metals and alloys: an advanced textbook in physical metallurgy. Pergamon Press, London
5. Garti N, Sato K (1988) Crystallization and polymorphism of fats and fatty acids. Marcel Dekker Inc., New York
6. Garti N, Sato K (2001) Crystallization processes in fats and lipid systems. Marcel Dekker Inc., New York
7. Marangoni AG (2012) Structure-function analysis of edible fats. AOCS Press, Urbana
8. Marangoni AG, Wesdorp LH (2013) Structure and properties of fat crystal networks, 2nd edn. CRC Press, Boca Raton

Chapter 8
Steady-State Nucleation Kinetics: The Fisher Turnbull Model

8.1 Nucleation

The nucleation process of crystalline material in foods structured by crystal networks is an important factor that strongly influences mechanical and functional properties, such as yield stress, hardness, texture and mouthfeel. For example, consider butter, although the same observations can be extended to fat materials such as palm oil, lard and tallow. Butter is a water-in-oil emulsion made from cream. Legally, butter must be composed of 80% milk fat. The remaining 20% is an aqueous solution containing protein, lactose and salt. In butter manufacture, the nucleation rate of the fat in the cream during the cream ageing process strongly influences the final texture of the butter. High supercooling conditions during aging favors the formation of numerous nuclei in the fat phase. The presence of numerous small crystals results in a stronger fat crystal network with smaller pores, which is very effective in entrapping liquid oil. During subsequent mechanical working of the solidified cream, the liquid oil thus remains entrapped within the fat globule instead of being squeezed out into the continuous phase. Liquid oil within the continuous phase serves as a lubricant, reducing the coefficient of friction between globules at stresses above the yield stress. Having less oil in the continuous phase thus results in a harder butter. On the other hand, slow crystallization favors crystal growth over the formation of nuclei. Large crystals are much less effective at trapping liquid oil. As a result, more liquid fat can be squeezed out into the continuous fat phase during working, resulting in a softer butter. The tribological properties of the liquid oil in the continuous phase are a key factor influencing the mechanical/textural response of butter.

A.G. Marangoni, *Kinetic Analysis of Food Systems*,
DOI 10.1007/978-3-319-51292-1_8

8.2 Energy of Activation for Nucleation (the Gibbs-Thomson Equation)

The formation of a surface during nucleus formation is an endergonic process as energy is required to counteract the surface tension between the growing solid phase and the existing supersaturated phase. The energy associated with the formation of a surface for a spherical nucleus is given by:

$$\Delta G_{\text{surface}} = 4\pi r^2 \delta \tag{8.1}$$

where $\Delta G_{\text{surface}}$ is free energy change associated with the formation of a surface, $4\pi r^2$ is the surface area of the crystal nucleus and δ is the solid-liquid surface energy or the crystal-melt interfacial tension (J·m^{-2}). As the solid-liquid surface energy remains constant, the energy required to maintain the surface is directly proportional to the surface area, which itself is a function of the radius of the crystal nucleus.

The creation of a nucleus, however, is not entirely an endergonic process. The creation of a nucleus also causes a decrease in the free energy of the system. This is driven by a decrease in the supersaturation of the system during crystallization. The expression relating the decrease in free energy to the decrease in supersaturation for a spherical nucleus is given by:

$$\Delta G_{\text{volume}} = -\frac{4}{3}\pi r^3 \frac{\Delta \mu}{V_m^s} \tag{8.2}$$

where ΔG_{volume} is the free energy change associated with a change in supersaturation, $\frac{4}{3}\pi r^3$ is the volume of the nucleus (m^3), $\Delta\mu$ is the difference in chemical potential between the liquid and the solid (J·mol^{-1}) and V_m^s is the volume of a mole of solid crystal (m^3·mol^{-1}).

As the molar volume V_m^s and the supersaturation $\Delta\mu$ remains relatively constant under a constant set of conditions, the ΔG_{volume} term is directly proportional to the volume of the nucleus, which is itself a function of the radius of the nucleus.

The balance between these two opposing processes thus determines whether or not a stable nucleus will form. The two free-energy terms can be combined to give the total free energy associated with the formation of a crystal nucleus. This expression is known as the Gibbs-Thomson equation. This formulation expresses the overall free energy change resulting from the formation of a spherical nucleus as the sum of the surface term and the volume term, namely:

$$\Delta G_n = \Delta G_{\text{surface}} + \Delta G_{\text{volume}} \tag{8.3}$$

which can be expanded into:

$$\Delta G_n = 4\pi r^2 \delta - \frac{4}{3}\pi r^3 \frac{\Delta \mu}{V_m^s} \tag{8.4}$$

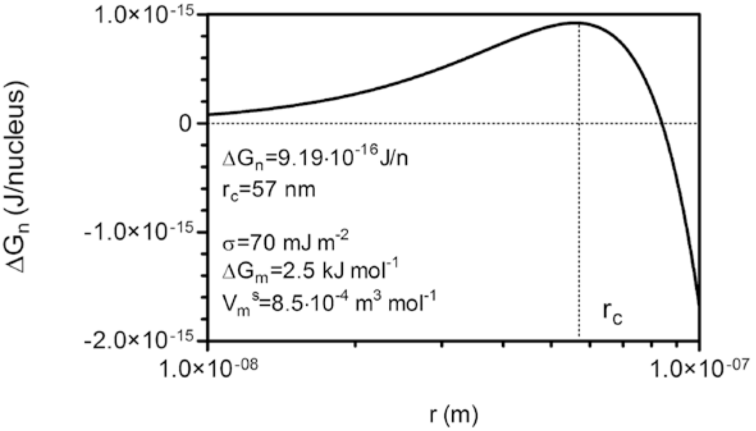


Fig. 8.1 Simulation of free energy changes as a function of nucleus radius. Notice the maximum in the energy profile which corresponds to the critical radius above which a stable nucleus will be formed

As previously mentioned, δ, $\Delta\mu$ and V_m^s can be regarded as constants for the purpose of this treatment. The free energy change associated with the formation of a spherical nucleus ΔG_n (in units of J) can thus be considered to be a function of solely the crystal radius r (in units of m). The balance between the exergonicity of nucleus formation (driven by the decrease in chemical potential between the melt and the crystal) and the endorgonicity due to the formation of a surface between the melt and the crystal) results in a maximum in the $\Delta G_n - r$ profile (Fig. 8.1). The crystal nucleus radius at the free energy maximum is the critical radius r_c. A stable nucleus forms when its radius is greater than the critical radius.

A radius below the critical radius implies a high surface area to volume ratio such that the energy required maintaining the surface exceeds that of the energy released by the formation of a crystal. Consequently, the free energy released upon nucleus formation is insufficient to overcome the free energy required to maintain the surface. At radii below the critical radius, the crystal nucleus will not form.

Envision a TAG component in a fat mixture. If the fat melt is cooled below the melting point of this component, it will crystallize until it reaches its saturation concentration in the melt. At this point, there is a thermodynamic equilibrium between the formation of crystals to dissolution of these crystals in the melt. This means that the chemical potential of the pure solid [$\mu_i^*(s)$] is equivalent to the chemical potential of that same component dissolved in the melt [$\mu_i(l)$], namely

$$\mu_i^*(s) = \mu_i(l) \tag{8.5}$$

The chemical potential of this particular component in the melt [$\mu_i(l)$] equals

$$\mu_i(l) = \mu_i^*(l) + RT\ln a_i(l) \tag{8.6}$$

where $\mu_i(l)$ is the chemical potential of a pure liquid of the component, and $a_i(l)$ corresponds to the activity of the TAG component in the melt, $a_i(l) = \gamma_i x_i(l)$, where γ_i is the activity coefficient and $x_i(l)$ is the concentration. Eq. 8.5 can be re-written as:

$$\mu_i^*(s) = \mu_i^*(l) + RT\ln a_i(l) \tag{8.7}$$

Re-arranging this equation gives the difference in the chemical potential between the solid and the melt as a function of the activity of the component in the melt:

$$\Delta\mu = \mu_i^*(s) - \mu_i^*(l) = +RT\ln a_i(l) \tag{8.8}$$

Consider the Hildebrand equation,

$$\ln a_i(l) = \frac{\Delta H_{f,i}}{R}\left(\frac{1}{T_{f,i}} - \frac{1}{T}\right) = \frac{\Delta H_{f,i}}{RT_{f,i}} - \frac{\Delta H_{f,i}}{RT} \tag{8.9}$$

where ΔH_f is the enthalpy of fusion (J·mol^{-1}), and T_f is the temperature of fusion (melting point).

Recall that $\Delta S_f = \dfrac{\Delta H_f}{T_f}$ for a process at equilibrium. By substituting this into the Hildebrand equation and multiplication with RT, the following form of the Hildebrand equation is obtained:

$$RT\ln a_i(l) = \frac{T\Delta H_{f,i}}{T_{f,i}} - \Delta H_{f,i} = T\Delta S_{f,i} - \Delta H_{f,i} = -\Delta G_{f,i} \tag{8.10}$$

Since $\Delta\mu = RT\ln a_i(l)$, Eq. 8.10 can be re-written as:

$$\Delta\mu = T\Delta S_{f,i} - \Delta H_{f,i} = -\Delta G_{f,i} \tag{8.11}$$

Thus, the difference in chemical potential between the pure solid of the TAG component and the pure liquid of the TAG component is equivalent to the free energy of fusion ΔG_f:

$$\Delta G_{f,i} = -RT\ln a_i(l) = \mu_i^*(l) - \mu_i^*(s) = -\Delta\mu \tag{8.12}$$

If the solid phase was heterogeneous with regards to the TAG composition, then the concentration (activity) of every component in the solid has to be taken into account. Thus, after the material has crystallized and equilibrium has been re-established, the chemical potential of the solid and the liquid would be equal,

$$\mu_i^*(s) + RT\ln a_i(s) = \mu_i^*(l) + RT\ln a_i(l) \tag{8.13}$$

The chemical potential difference between the pure TAG component solid and the pure TAG component liquid could then be expressed as a function of the concentration (activity) of the specific component in the solid and the liquid,

$$\begin{aligned}\Delta\mu &= \mu_i^*(s) - \mu_i^*(l) = RT\ln a_i(l) - RT\ln a_i(s)\\ &= RT\ln\frac{a_i(l)}{a_i(s)}\end{aligned} \tag{8.14}$$

Considering the above, we can rewrite the Gibbs-Thomson equation as:

$$\Delta G_n = 4\pi r^2\delta - \frac{4}{3}\pi r^3\frac{\Delta G_f}{V_m^s} \tag{8.15}$$

The change in free energy (derivative of the Gibbs-Thomson equation) as a function of radius is given by:

$$\frac{\partial\Delta G_n}{\partial r} = 8\pi r\delta - 4\pi r^2\frac{\Delta G_f}{V_m^s} \tag{8.16}$$

The formation of a stable crystal nucleus takes place above a critical radius (r_c), that is the energetics of the system allow for the formation of nuclei with radii above this critical radius r_c. This radius corresponds to the maximum in the free energy versus nuclei radius profile. At this point, the change in free energy relative to the radius r is zero:

$$\frac{\partial\Delta G_n}{\partial r} = 0 \tag{8.17}$$

Equation 8.16 can then be rearranged to

$$8\pi r_c\delta = 4\pi r_c^2\frac{\Delta G_f}{V_m^s} \tag{8.18}$$

Re-arranging Eq. 8.18 gives the following expression for the critical radius:

$$r_c = \frac{2\delta V_m^s}{\Delta G_f} \tag{8.19}$$

Substituting the expression for r_c (Eq. 8.19) for r in the Gibbs-Thompson equation (Eq. 8.4) gives:

$$\Delta G_n^{r_c} = 4\pi \left(\frac{2\delta V_m^s}{\Delta G_f} \right)^2 \delta - \frac{4}{3}\pi \left(\frac{2\delta V_m^s}{\Delta G_f} \right)^3 \frac{\Delta G_f}{V_m^s} \tag{8.20}$$

which, after rearrangement, leads to the following useful expression:

$$\begin{aligned} \Delta G_n^{r_c} &= \frac{16\pi\delta^3 \left(V_m^s\right)^2}{\left(\Delta G_f\right)^2} - \frac{32\pi\delta^3 \left(V_m^s\right)^3 \Delta G_f}{3\left(\Delta G_f\right)^3 V_m^s} = \frac{16\pi\delta^3 \left(V_m^s\right)^2}{\left(\Delta G_f\right)^2} - \frac{32\pi\delta^3 \left(V_m^s\right)^2}{3\left(\Delta G_f\right)^2} \\ &= \frac{48\pi\delta^3 \left(V_m^s\right)^2}{3\left(\Delta G_f\right)^2} - \frac{32\pi\delta^3 \left(V_m^s\right)^2}{3\left(\Delta G_f\right)^2} \\ &= \frac{16\pi\delta^3 \left(V_m^s\right)^2}{3\left(\Delta G_f\right)^2} \end{aligned} \tag{8.21}$$

Considering that

$$\Delta G_f = \Delta H_f - T\Delta S_f \tag{8.22}$$

and that since $\Delta G_f = 0$ when $T = T_f$, then

$$\Delta S_f = \frac{\Delta H_f}{T_f} \tag{8.23}$$

The free energy of fusion can be thus expressed as:

$$\begin{aligned} \Delta G_f &= \Delta H_f - T\left(\frac{\Delta H_f}{T_f} \right) = \Delta H_f \left(1 - \frac{T}{T_f} \right) = \Delta H_f \left(\frac{T - T_f}{T_f} \right) \\ &= \frac{\Delta H_f \left(\Delta T \right)}{T_f} \end{aligned} \tag{8.24}$$

Substituting Eq. 8.24 into Eq. 8.21 results in the common form of the Gibbs-Thompson equation used in crystallization studies:

$$\Delta G_n^{r_c} = \frac{16\pi\delta^3 \left(V_m^s\right)^2 T_f^2}{3\Delta H_f^2 \left(\Delta T\right)^2} \tag{8.25}$$

This free energy of nucleation corresponds to an activation free energy of nucleation – the free energy barrier required to form a nucleus of critical radius r_c.

8.3 The Fisher-Turnbull Equation

The Fisher-Turnbull equation can be used to quantify the activation free energy of nucleation for a crystallization process. The Fisher-Turnbull equation can be obtained by considering the activation free energy of nucleation to be a combination of a free energy of activation for diffusion ($\Delta G_d^{\#}$) and the free energy of nucleation for a nucleus of critical radius r_c ($\Delta G_n^{r_c}$, Eq. 8.25):

$$\Delta G^{\#} = \Delta G_d^{\#} + \Delta G_n^{r_c} \tag{8.26}$$

The derivation of the Fisher-Turnbull model involves concepts derived in the absolute reaction rate theory, where the rate of a first-order reaction can be written as a function of its activation energy. For the sake of clarity, the reaction rate v can be written as follows:

$$v = \frac{Nk_B T}{h} e^{-\frac{\Delta G^{\#}}{RT}} \tag{8.27}$$

Substituting Eq. 8.26 into Eq. 8.27 gives the following expression for the rate of nucleation:

$$v = \frac{Nk_B T}{h} e^{-\frac{\Delta G_d^{\#}}{kT}} e^{-\frac{\Delta G_n^{r_c}}{kT}} \tag{8.28}$$

where v now denotes the rate of nucleation (s^{-1}), and N the number of molecules in the melt participating in the nucleation process.

$\Delta G_d^{\#}$ is assumed to remain constant at relatively low degrees of undercooling. Likewise, for a narrow range of temperatures, $\Delta G_d^{\#}$ can also be assumed to remain constant. Taking this into consideration, Eq. 8.28 can be expressed as:

$$v = \frac{\alpha Nk_B T}{h} e^{-\frac{\Delta G_n^{r^*}}{kT}} \tag{8.29}$$

where the contribution of this diffusion term is assumed to be a constant α:

$$\alpha = e^{-\frac{\Delta G_d^{\#}}{kT}} \tag{8.30}$$

Experimentally, the rate of nucleation is determined by taking the inverse of the induction time of crystallization (τ):

$$v \sim \frac{1}{\tau} \tag{8.31}$$

Equation 8.29 can thus be expressed as:

$$\frac{1}{\tau} \sim \frac{\alpha N k_B T}{h} e^{-\frac{\Delta G_n^{r_c}}{kT}} \tag{8.32}$$

which can be re-arranged to give:

$$\frac{1}{\tau T} \sim \frac{\alpha N k_B}{h} e^{-\frac{\Delta G_n^{r_c}}{kT}} \tag{8.33}$$

Taking the natural logarithm on both sides results in:

$$\ln\left(\tau T\right) \sim -\ln\frac{\alpha N k_B}{h} + \frac{\Delta G_n^{r_c}}{kT} \tag{8.34}$$

Introducing Eq. 8.25 into Eq. 8.34 yields:

$$\ln\left(\tau T\right) \sim -\ln\frac{\alpha N k_B}{h} + \frac{\Delta G_n^{r_c}}{kT} = -\ln\frac{\alpha N k_B}{h} + \frac{\left(\frac{16\pi\delta^3\left(V_m^s\right)^2 T_f^2}{3\Delta H_f^2\left(\Delta T\right)^2}\right)}{kT} \tag{8.35}$$

Subsequent rearrangement of this equation gives:

$$\ln\left(\tau T\right) \sim -\ln\frac{\alpha N k_B}{h} + \frac{16\pi\delta^3\left(V_m^s\right)^2 T_f^2}{3k\Delta H_f^2}\left(\frac{1}{T\left(\Delta T\right)^2}\right) \tag{8.36}$$

Thus, a plot of ln(τT) versus $\frac{1}{\left(T\left(\Delta T\right)^2\right)}$ should yield a straight line with a slope m (K^3), which equals:

$$m = \frac{16\pi\delta^3\left(V_m^s\right)^2 T_f^2}{3k_B\Delta H_f^2} \tag{8.37}$$

Since

$$\Delta G_n^{r_c} = \frac{16\pi\delta^3\left(V_m^s\right)^2 T_f^2}{3\Delta H_f^2\left(\Delta T\right)^2} \tag{8.25}$$

the free energy of nucleation can be determined from:

$$\Delta G_n^{r_c} = \frac{m k_B}{\left(\Delta T\right)^2} \tag{8.38}$$

Figure 8.2 depicts the Fisher-Turnbull plot for a blend of 12% Palsgaard 6111 stabilizer (42% behenic, 42% stearic, 5% palmitic acid) in high-oleic safflower oil. The graph displays a single line. This is indicative of a simple system where only one component is crystallizing at a given time. For more complex and multicomponent systems such as palm oil, several different components may crystallize over a given temperature and supersaturation range. The result is a graph consisting of different lines. These lines can be analyzed individually and the $\Delta G_n^{r_c}$ determined for each independent crystallization process.

In our case, the peak melting temperature of the 12% blend of Palsgaard 6111 in high oleic safflower oil is 47 °C, as determined by differential scanning calorimetry. The system is thus supersaturated below 47 °C. The slope of this line can be obtained and the activation energy of nucleation calculated. The plot of the nucleation free energy against temperature is given in Fig. 8.3. As the set crystallization temperature T increases and approaches the melting point T_f, the supersaturation ΔT decreases. Likewise, as T increases, the free energy of nucleation increases. Thus, a temperature increase such that T approaches the melting point will result in a decreased rate of nucleation.

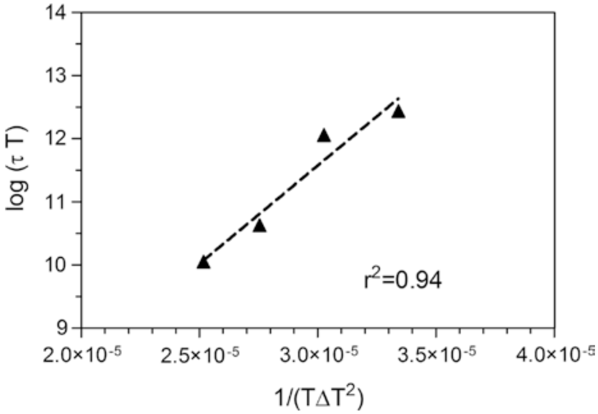


Fig. 8.2 Fisher-Turnbull plot for a 12% Palsgaard 6111 in high oleic safflower oil created using induction times of nucleation in the range 40–50 °C. Induction times of nucleation (τ, sec) were determined using a cloud point analyzer (turbidity measurements)

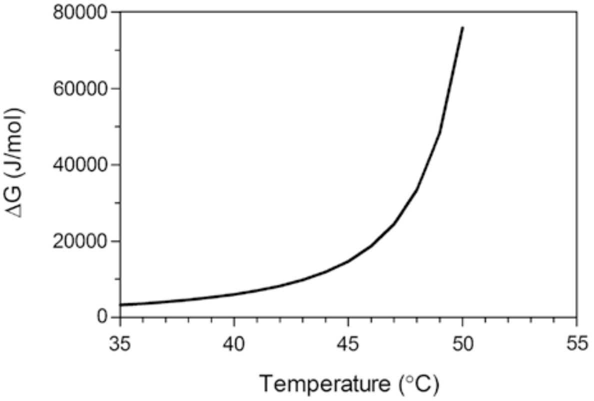


Fig. 8.3 Changes in the activation free energies of nucleation (J/mol) as a function of temperature (°C) for 12% Palsgaard 6111 in high oleic safflower oil, calculated from the slope obtained in Fig. 8.2

8.4 Other Applications of the Fisher-Turnbull Model

Another useful application of the Fisher-Turnbull equation is in the determination of the crystal-melt interfacial tension (δ). If reliable calorimetric data for the enthalpy and temperature of fusion are available, and an estimate of the molar volume can be obtained, the solid-melt interfacial tension can be directly calculated from:

$$\delta = \left(\frac{3mk_B \Delta H_f^2}{16\pi \left(V_m^s \right)^2 T_f^2} \right)^{\frac{1}{3}} \tag{8.39}$$

Estimates of the enthalpy of fusion should be obtained experimentally using differential scanning calorimetry. The density of a solid TAG mixture, crystallized under relevant conditions, should also be determined experimentally by picnometry. The molar volume of a fat can then be calculated from density data using

$$molar\ \ volume\ \left(m^3 / mol\right) = \frac{MW\left(g / mol\right)}{density\left(g / ml\right)} \left(1 \cdot 10^{-6} m^3 / ml\right) \tag{8.40}$$

Bibliography

1. Christian JW (1965) The theory of transformations in metals and alloys: an advanced textbook in physical metallurgy. Pergamon Press, London
2. Garti N, Sato K (1988) Crystallization and polymorphism of fats and fatty acids. Marcel Dekker Inc., New York
3. Garti N, Sato K (2001) Crystallization processes in fats and lipid systems. Marcel Dekker Inc., New York
4. Marangoni AG (2012) Structure-function analysis of edible fats. AOCS Press, Urbana
5. Marangoni AG, Wesdorp LH (2013) Structure and properties of fat crystal networks, 2nd edn. CRC Press, Boca Raton

Chapter 9
Non-isothermal Nucleation Kinetics in Fats

9.1 Non-isothermal Nucleation

The model discussed in Chap. 8 is for a special case of the nucleation phenomenon. The isothermal model addresses situations where the temperature remains constant during the reaction and the temperature drop from the melting temperature, T_m, to the set crystallization temperature, T_{set}, is instantaneous (Fig. 9.1a). As well, it is assumed that crystallization occurs only when T_{set} is reached and not prior to this. For the isothermal case, time zero is assumed to be the start of the experiment as the temperature at the beginning of the experiment is assumed to be T_{set}. This model is suited to the study of systems that are not heat transfer limited.

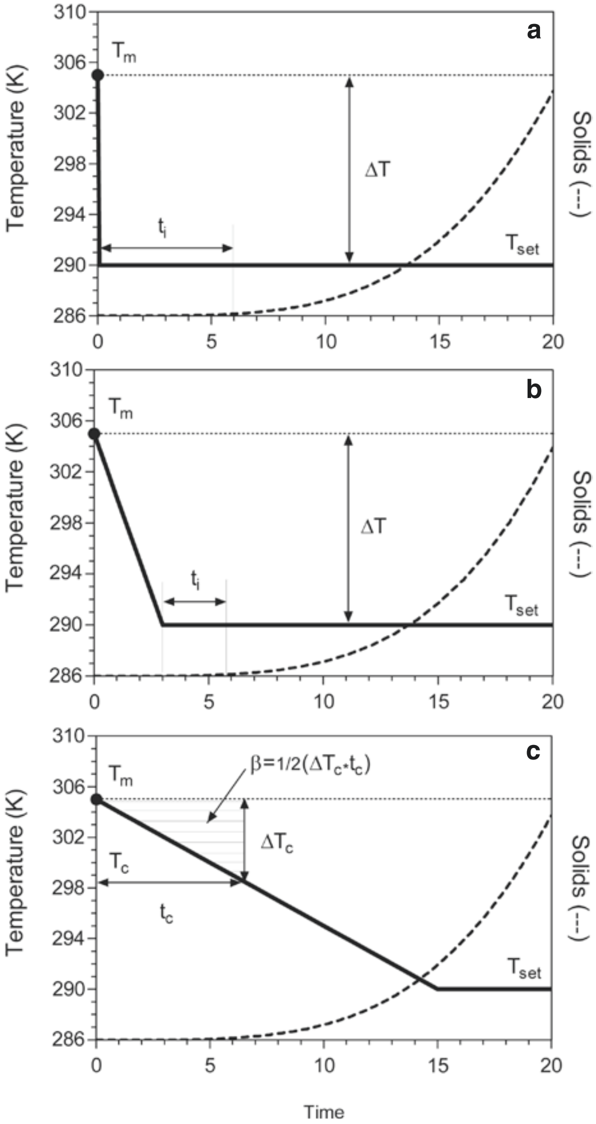
Under isothermal conditions, the crystallization process can be characterized by an induction time (t_i), which is the time required for the appearance of the first solid nuclei at T_{set} under the influence of a thermal driving force. The induction time t_i is proportional to the degree of supersaturation (solutions), or degree of supercooling or undercooling (melts), ΔT, which is the difference between the equilibrium melting temperature of the material (T_m) and the set crystallization temperature (T_{set}), $T_m - T_{set}$.

Experimental realities limit the speed at which a system can reach the set temperature. Limitations in heat transfer will result in a gradual reduction in temperature (Fig. 9.1b) as opposed to an instantaneous drop observed in Fig. 9.1a. These conditions can be considered "near-isothermal". It is assumed, in this case, that crystallization does not begin until after some time the set crystallization temperature has been reached. For a "near-isothermal" case, then, it is still possible to determine an induction time of nucleation and treat the crystallization process as if it was taking place under isothermal conditions. For this case, it is important to remember that time zero corresponds to the time when the system reaches T_{set}. The induction time in this case, is the time interval between the attainment of the set crystallization temperature and the time of the first appearance of solid nuclei.

In the industrial manufacture of edible fat products, such as margarines and chocolate, from multicomponent triacylglycerol mixtures, crystallization takes

A.G. Marangoni, *Kinetic Analysis of Food Systems*,
DOI 10.1007/978-3-319-51292-1_9

Fig. 9.1 Temperature and crystallization profiles for materials crystallizing under isothermal (**a**), near isothermal (**b**) and non-isothermal (**c**) conditions. Reproduc



place under non-isothermal conditions. Under these conditions, crystallization occurs prior to attaining the set crystallization temperature T_{set} (Fig. 9.1c). For this situation, both the non-isothermal induction time of nucleation (t_c) and the undercooling at nucleation (ΔT_c) have different meanings than for the isothermal case.

The crystallization behavior and structure of these materials under such conditions is extremely sensitive to heat and mass transfer conditions. The crystallization regime will ultimately affect mechanical strength, flow behavior and sensory texture. Of particular interest is the nucleation behavior of such systems as important structural features, such as crystallite number, size and morphology, as well as the spatial distribution of mass, are a direct consequence of nucleation behavior. However, no good theoretical tools exist to model the nucleation behavior of these complex organic mixtures under non-isothermal conditions. In this section, a new approach to the modeling of non-isothermal nucleation systems will be presented.

9.2 Formulation of the Time-dependent Supercooling Parameter

To model non-isothermal nucleation, a new quantity or parameter that defines the driving force for nucleation has to be defined. This parameter will be shown to embody the dynamic of the system. It was previously mentioned that the induction time of nucleation and undercooling at nucleation have different meanings for the isothermal vs. non-isothermal cases. The first step, therefore, in the formulation of a non-isothermal model, is the re-definition of these parameters. Upon examination of Fig. 9.1c, one can see that it is not the temperature differential, ΔT, that is the driving force for nucleation. Instead, it is the time the system has been exposed to a particular temperature differential. In other words, the supercooling of the system is a dynamic quantity as opposed to being a static quantity as for the isothermal case: ΔT is changing in time as the material crystallizes and changes composition. Thus, a supercooling-time exposure has to be defined. This corresponds to the exposure of the system to supercooling until the initiation of nucleation. This is calculated as the area under the supercooling-time trajectory from the time when the system crosses the melting temperature, T_m to the time where the first crystal nuclei appear (t_c). The temperature at which the first crystal nuclei appear is called the crystallization temperature,T_c. Notice that $T_c \neq T_{set}$ and therefore $[\Delta T = (T_m - T_{set})] \neq [\Delta T_c = (T_m - T_c)]$. If the cooling rate is assumed to be constant, the supercooling-time exposure (β^2) at the onset of nucleation can thus be defined as:

$$\beta^2 = \frac{1}{2} \Delta T_c \cdot t_c \tag{9.1}$$

Another important parameter in the characterization of a non-isothermal system is the cooling rate. A linear cooling rate is defined as:

$$\phi = \frac{\Delta T}{\Delta t} \tag{9.2}$$

As previously discussed, at the crystallization temperature, T_c, $t = t_c$. At the melting temperature, T_m, $t_0 = 0$. Substituting these into Eq. 9.2 gives:

$$\phi = \frac{(T_m - T_c)}{t_c - 0} = \frac{\Delta T_c}{t_c} \tag{9.3}$$

Substituting $\frac{\Delta T_c}{\phi}$ for t_c in the expression for the supercooling-time exposure at nucleation (Eq. 9.1) gives the following expression:

$$\beta^2 = \frac{1}{2}\frac{\left(\Delta T_c\right)^2}{\phi} \tag{9.4}$$

The parameter β^2 corresponds to the triangular area under the curve for the supercooling-time curve (Fig. 9.1c). This parameter takes into consideration the amount of supercooling in time required for nucleation to start. It is important to realize that the parametrization of the data relative to temperature as well as time is necessary for a proper description of nucleation under non-isothermal conditions. The time-dependent supercooling parameter β^2 incorporates a thermodynamic component in the form of the supercooling at nucleation (ΔT_c) *and* a kinetic component in the form of the nonisothermal nucleation induction time (t_c).

Finally, it is the square root of beta corresponds then to an effective supercooling experienced by the system at nucleation:

$$\beta = \frac{\Delta T_c}{\sqrt{2\phi}} \tag{9.5}$$

9.3 A Probabilistic Approach to Modeling Non-isothermal Nucleation Kinetics

The normalized nucleation rate ($\frac{J}{J_{\max}}$) was found to have an exponential dependence on β, where $J_{\max}$ is the maximum nucleation rate. The exponential relationship between the normalized nucleation rate and the effective supercooling parameter is shown in Fig. 9.2. Figure 9.3 shows the dependence of the normalized nucleation rate on the cooling rate.

The shape observed in Fig. 9.2 led to the conclusion that perhaps non-isothermal nucleation kinetics could be modeled statistically, in a similar fashion as for the kinetic theory of gases. In order to develop this argument logically, we must revisit the basic premises of kinetic theory.

The rate of a reaction (v) is a function of the concentration of molecules with sufficient energy to overcome an energy barrier to the particular reaction (N^*), and thus $v = k[N^*]$, where k is the rate constant for the reaction and N^* corresponds to the concentration of molecules in the activated state. In the kinetic theory of gases, the molecules in the activated state are those molecules with sufficient energy, and in the proper orientation, to undergo the chemical reaction. For the case of nucleation reactions, N^* would correspond to the concentration of molecules in the metastable

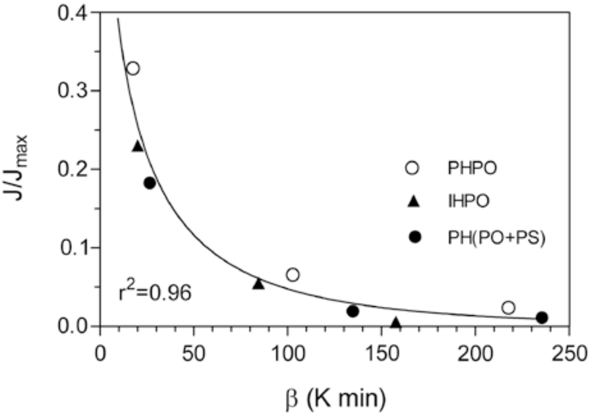


Fig. 9.2 Dependence of the relative nucleation rate on the supercooling-time exposure, namely the effective supercooling (β) in partially hydrogenated palm oil (*PHPO*), interesterified, hydrogenated palm oil (*IHPO*) and a partially hydrogenated mixture of palm oil and palm stearin PH(*PO*+*PS*)

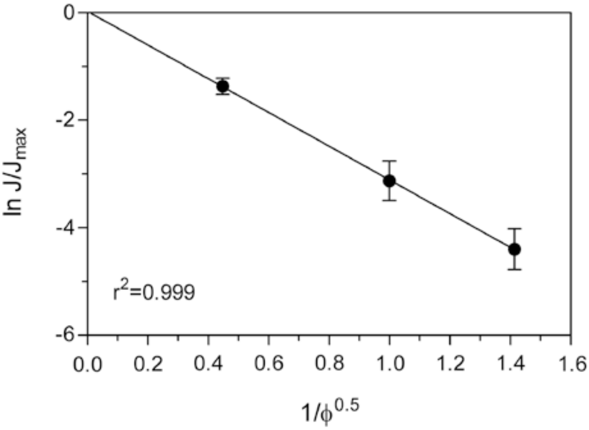


Fig. 9.3 Dependence of the relative nucleation rate on cooling rate for partially hydrogenated palm oil (*PHPO*), interesterified, hydrogenated palm oil (*IHPO*) and a partially hydrogenated mixture of palm oil and palm stearin PH(*PO*+*PS*)

state, just prior to the nucleation event. The proportion of molecules in the appropriate state to undergo a reaction (from energetic and conformational considerations) will be given by $(N^*) = p(x)(N_T)$, where N_T is the total concentration of reactant molecules, and p(x) corresponds to the probability density function (*pdf*) which describes the frequency distribution of the particular event.

Following this line of reasoning, the effective supercooling-time parameter β was assumed to be distributed in an exponential fashion, with an exponential *pdf*, $p(\beta; k)$, of the form:

$$p(\beta;k) = \begin{Bmatrix} ke^{-k\beta}; \beta \geq 0 \\ 0 \quad\quad ; \beta < 0 \end{Bmatrix} \tag{9.6}$$

The parameter k is called the rate parameter. The rate parameter has to satisfy the condition $k > 0$. This *pdf* applies to values of the randomly distributed variable belonging to the set $\beta \in [0; \infty)$. The scale parameter (μ) is simply the inverse of the rate parameter and represents the mean, or expected value, of an exponentially distributed random variable, $E\left[\beta\right] = \mu = \frac{1}{k}$. Thus, this *pdf* is appropriate to model our situation where our random variable has to always be greater than zero, and the mean is fixed.

Exponential distributions are used to model memory-less Poisson, or stochastic processes, which take place with constant probability per unit time or distance. This is the reason why exponential *pdfs* are extensively used to model random processes such as Brownian motion. In the case of Brownian motion, the future position of a molecule is independent of its current position. In our case, we assume that our nucleation phase transition initiation event takes place with a constant probability per unit effective supercooling in time (β), possibly not unreasonable considering the constant cooling rates used.

Another interesting property of an exponential *pdf* is that among all continuous pdfs, with support [0; ∞), the exponential pdf with $\mu = 1/k$ has the highest entropy. Many physical systems tend to move towards maximal entropy configurations over time (Principle of Maximum Entropy).

Considering all of the above, the rate of the nucleation reaction (J) will thus be given by:

$$J = k_p N_T k e^{-k\beta} = J_{\max} k e^{-k\beta} \tag{9.7}$$

9.4 Determining the Energy of Activation for a Non-isothermal Process

By combing Eqs. 9.5 and 9.7, the relative nucleation rate can be expressed as an exponential function of the inverse of the square root of the cooling rate, namely:

$$\frac{J}{J_{\max}} = k e^{-k\frac{\Delta T_c}{\sqrt{2\phi}}} \tag{9.8}$$

Using this model, it is possible to determine the energy required to initiate the nucleation process. Consider that from T_m to T_c, no phase change has taken place. Thus, up to this point, strictly specific heat Q_m has been removed from the system ($Q_m = C_p \Delta T$), where Q_m is the specific heat removed from the system upon cooling per gram of material, C_p is the specific heat (heat capacity) of the material. Substituting $\frac{Q_m}{C_p}$ for ΔT in Eq. 9.8 leads to the expression:

$$\frac{J}{J_{\max}} = k e^{\frac{-kQ_m}{C_p\sqrt{2\phi}}} = k e^{\frac{-Q_m}{Z\sqrt{\phi}}} = k e^{\frac{-X}{\sqrt{\phi}}} \tag{9.9}$$

where Q_m is proposed here to represent the energy of activation for nucleation per unit mass [J/g], and Z [J g^{-1} K$^{-1/2}$ s$^{1/2}$] is defined as:

$$Z = \frac{\sqrt{2}C_p}{k} \tag{9.10}$$

Thus, from knowledge of k (Eq. 9.7) and C_p, Z can be calculated using Eq. 9.10. Moreover, from the nonlinear fit of $J/J_{\max}$ vs $1/\sqrt{\phi}$ data, the constant X can be obtained. It is then possible to determine the energy of activation for nucleation as $Q_m = Z * X$ [J/g]. This quantity can then be multiplied by the average molecular weight of the triacylglycerols [MW, g/mol], to obtain the molar energy of activation for nucleation (Q_m), $Q_M = Z * X * MW$ [J/mol].

9.5 Special Case When β Is Very Small

As shown in Eq. 9.7, the rate of nucleation for a non-isothermal crystallization process can have the form:

$$\frac{J}{J_{\max}} = ke^{-k\beta} \tag{9.11}$$

The first two terms of the Taylor expansion of this function are

$$\frac{J}{J_{\max}} = ke^{-ka} - k^2 e^{-ka}\left(\beta - a\right) \tag{9.12}$$

For the case where $a = 0$, this expression reduces to

$$\frac{J}{J_{\max}} \approx k - k^2\beta \tag{9.13}$$

Thus, for the case where the time-dependent supercooling parameter (β) is close to 0, the exponential probability density function can be approximated by a linear function. This implies that such type of nucleation process has a very small metastable region and therefore initiated at very small supersaturations ($\Delta T \approx 0$), and/or with very short induction times ($t_i \approx 0$).

In order to determine the energy of activation using this expression, we can use the same arguments as above. Replacing β with $\frac{\Delta T_c}{\sqrt{2\phi}}$ and ΔT_c with Q_m/C_p, we obtain

$$\frac{J}{J_{\max}} = k - \frac{k^2\Delta T_c}{\sqrt{2\phi}} = k - \frac{k^2 Q_m}{C_p\sqrt{2\phi}} \tag{9.14}$$

If parameters are grouped as for Eq. 9.10, $Z = \frac{\sqrt{2}C_p}{k}$ and $X = \frac{Q_m}{Z}$, then Eq. 9.14 can be expressed as:

$$\frac{J}{J_{\max}} = k - \frac{kQ_m}{Z\sqrt{\phi}} = k - \frac{kX}{\sqrt{\phi}} \tag{9.15}$$

Thus, a plot of J/J_{max} vs. $1/\sqrt{\phi}$ will yield a straight line with y-intercept k, and slope kX. Using the obtained parameter k, Z can be calculated using $Z = \frac{\sqrt{2}C_p}{k}$, while Q_m can then be determined from $Q_m = X * Z$.

9.6 Determination of the Induction Time

The following section describes the application of the proposed model in the study of non-isothermal nucleation kinetics in materials such as anhydrous milk fat (AMF), partially hydrogenated palm oil (PHPO), palm oil (PO), chemically interesterified and hydrogenated palm oil (IHPO), and a partially hydrogenated blend of palm oil and palm stearin (PH(PO+PS)). The temperature T_c and time t_c for the appearance of the first solid crystals were determined using a Phase Transition Analyzer (PTA). This machine is essentially a very sensitive turbidimeter that detects a phase transition during a crystallization process. It incorporates a crystallization cell to cool the sample at a constant cooling rate. Melted samples are injected into this cell. Upon injection, the sample is cooled at a constant cooling rate. Both the crystallization temperatures and the induction times were determined from these experiments as the temperature and time when the first crystals appear, which is evidenced by a deviation in the signal baseline. The induction time for non-isothermal nucleation (t_c) was calculated as the time at which the first crystals were detected minus the time required to reach the melting point temperature.

9.7 Determination of the Melting Point

The PTA technique was also used to determine the melting points of the fats. Melted samples (80 °C for 30 minutes) were placed in the PTA cell at 60 °C and then cooled to 10 °C at 40 °C/min. Upon reaching this temperature, samples were incubated at this temperature for 30 min and then heated at 1 and 5 °C/min until melted. The melting temperature was determined as the temperature at which the signal of the laser becomes constant and equal to the baseline. At this point, all the solids have melted away as evidenced by the lack of turbidity in the sample.

Determination of the melting point of a pure crystalline solid is straightforward and unambiguous. It will melt at a specific temperature regardless of the method used to determine the melting point. It must be remembered, however, that fats are heterogeneous multicomponent materials, that is, they are almost always composed of several different components. It is difficult to determine a specific melting point for a fat as the melting point is influenced by a number of factors, not least of all the presence of different polymorphic forms in the solid state. In fact, fats do not have a true melting point unlike relatively pure materials. The thermal properties of a fat are instead characterized by a melting range. Depending on the technique used and the thermal history of the sample, different melting points can be obtained. For this reason, it is very important to specify which techniques and crystallization conditions were used to determine the 'melting point'.

In the analysis carried out in this work, an unabmiguous melting temperature was needed for the purpose of kinetic analysis. This represented a challenge since this temperature has to be related to the initial triacylglycerol fraction that nucleates. Moreover, the determination of this average/global melting temperature had to be reasonably reproducible.

Figure 9.4 shows the melting profile of the 5 fat samples obtained using the PTA. These melting profiles will be used to determine a global average melting point for the fat being observed. The melting temperature is determined here as the temperature at which the signal intensity decreases to baseline and becomes constant. Interestingly, this method is equivalent to the commonly used Capillary Melting Point determination method, which relies on a visual inspection of a capillary to determine at which point the material melts and becomes transparent. The melting points of the different fats are given in Table 9.1.

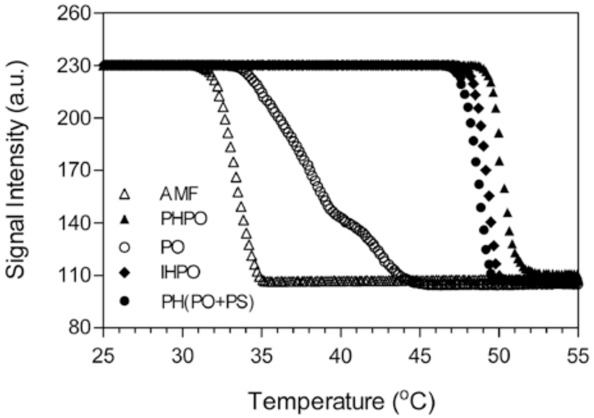


Fig. 9.4 Phase transition analyzer (light scattering) profiles as a function of increasing temperature used to determine the melting point of the different fat systems: anhydrous milkfat (*AMF*), palm oil (*PO*), partially hydrogenated palm oil (*PHPO*), interesterified, hydrogenated palm oil (*IHPO*) and a partially hydrogenated mixture of palm oil and palm stearin PH(*PO*+*PS*)

Table 9.1 The global melting points of five different fat samples obtained by determining the temperature at which the signal intensity in the PTA reaches baseline

Fat sample	Melting point (°C)
IHPO	49.8 ± 0.9
AMF	35.5 ± 0.2
PH(PO/PS)	49.9 ± 0.2
PHPO	51.3 ± 1.7
PO	45.6 ± 0.4

IHPO interesterified, hydrogenated palm oil, *AMF* anhydrous milkfat, *PH(PO+PS)* partially hydrogenated mixture of palm oil and palm stearin, *PHPO* partially hydrogenated palm oil, *PO* palm oil
Values represent the average and standard deviation of three replicate samples

This method enables the determination of melting points as the temperature when the last crystals in the fat melt away. This method offers a convenient way to determine an average value for the melting point of a fat. Certainly this is not the true melting point of the fat, since the fat is composed of hundreds of different triacyglycerol molecules, each with a unique melting point. The practice of using a global melting temperature was only resorted to in order to be able to carry out the analysis. Otherwise, the analysis would not be possible.

Table 9.2 shows the melting and crystallization temperatures obtained from differential scanning calorimetry (DSC) and PTA together with the induction times of crystallization determined by PTA. The DSC peak melting temperatures and the PTA end of melt temperatures reported in this table were determined from the melting profiles obtained after incubating the samples for 5 min (DSC) and 30 min (PTA) after reaching T_{set}. Thus, they represent the melting temperature of the solids which are formed at the onset of the crystallization process.

9.8 Determination of the Nucleation Rate

The nucleation rate was estimated from the inverse of the induction time of nucleation obtained from the PTA experiments as given in Table 9.2. This method is an approximation, but it is experimentally accessible and convenient. A nucleation rate corresponds to the number of nuclei appearing per unit time, however, here we approximate this rate from the inverse of an induction time of nucleation. It would be prudent, therefore, to compare this method to a more direct, but less convenient, method of determining nucleation rate, such as by light microscopy. The nucleation rate was therefore determined using polarized light microscopy by counting the number of particles appearing per unit time. Figure 9.5a–c show the evolution of the number of crystals appearing as a function of time in three fat systems at three cooling rates, and the corresponding first derivatives of the N vs. t curves are shown in panels 9.5d–f. The peak value of the first derivative was taken as the nucleation rate (J_p). As can be appreciated in Fig. 9.6, increasing the cooling rate leads to an increase

Table 9.2 Peak melting temperatures (T_m), and peak crystallization temperatures (T_c) of samples cooled at as different rates (ϕ) determined using differential scanning calorimetry (DSC), melting (T_m) and crystallisation (T_c) temperatures determined by phase transition analysis (PTA), and induction time of nucleation (T_c) determined using PTA

ϕ (°C/min)	T_m (DSC)(°C)	T_c (DSC)(°C)	T_m (PTA)(°C)	T_c (PTA)(°C)	t_c (s)
AMF					
0.5	33.6 ± 0.1	17.8 ± 0.7	35.5 ± 0.2	21.4 ± 0.6	1734 ± 69
1	33.6 ± 0.2	16.2 ± 0.2		19.9 ± 0.4	956 ± 25
2	34.0 ± 0.2	15.1 ± 0.1		18.9 ± 0.3	508 ± 10
3	34.3 ± 0.1	15.6 ± 1.4		18.8 ± 0.2	343 ± 14
4	34.2 ± 0.3	14.7 ± 0.2		18.3 ± 0.5	264 ± 7
5	33.5 ± 0.1	14.4 ± 0.1		19.5 ± 0.1	195 ± 22
10				16.0 ± 0.2	119 ± 6
15				15.8 ± 0.9	81 ± 7
20				15.9 ± 1.0	60 ± 5
PHPO					
0.5	45.7 ± 0.4	27.3 ± 0.1	51.3 ± 1.7	34.6 ± 0.1	2050 ± 14
1	45.7 ± 0.1	27.7 ± 0.2		32.9 ± 1.7	1128 ± 106
2	46.0 ± 0.2	27.4 ± 0.1		30.6 ± 0.1	635 ± 8
3	45.7 ± 0.1	27.3 ± 0.1		30.4 ± 0.3	428 ± 8
4	45.6 ± 0.3	27.2 ± 0.1		30.1 ± 0.5	326 ± 8
5	45.6 ± 0.1	27.1 ± 0.1		31.2 ± 0.7	247 ± 10
10				29.0 ± 0.9	136 ± 7
15				28.4 ± 0.6	94 ± 12
20				28.7 ± 0.8	70 ± 3
PO					
0.5	43.7 ± 0.1	20.3 ± 0.01	45.6 ± 0.4	25.3 ± 1.0	2471 ± 169
1	43.0 ± 0.1	19.8 ± 0.1		24.0 ± 0.1	1323 ± 18
2	43.3 ± 0.1	19.5 ± 0.1		23.0 ± 0.3	691 ± 10
3	43.0 ± 0.3	19.5 ± 0.1		22.9 ± 0.8	466 ± 16
4	43.0 ± 0.3	19.4 ± 0.1		21.7 ± 0.8	359 ± 21
5	42.8 ± 0.5	19.4 ± 0.1		23.0 ± 0.8	281 ± 13
10				21.3 ± 0.3	149 ± 8
15				21.1 ± 1.5	100 ± 7
20				21.1 ± 1.2	75 ± 16
IHPO					
0.5	47.2 ± 0.5	31.9 ± 1.8	49.8 ± 0.9	37.4 ± 0.4	1527 ± 57
1	46.7 ± 0.1	32.0 ± 0.1		34.5 ± 0.2	937 ± 10
2	47.0 ± 0.5	31.3 ± 0.1		34.1 ± 0.2	483 ± 6
3	46.6 ± 0.4	29.8 ± 0.1		33.9 ± 0.1	325 ± 4
4	32.7 ± 0.6	30.1 ± 0.1		33.8 ± 0.3	230 ± 66
5	33.2 ± 0.2	30.1 ± 0.2		34.6 ± 1.1	187 ± 16
10				33.0 ± 0.3	103 ± 9
15				32.6 ± 0.7	70 ± 12

(continued)

Table 9.2 (continued)

ϕ (°C/min)	T_m (DSC)(°C)	T_c (DSC)(°C)	T_m (PTA)(°C)	T_c (PTA)(°C)	t_c (s)
20				32.4 ± 0.2	54 ± 16
PH (PO+PS)					
0.5	45.7 ± 0.3	30.0 ± 0.1	49.9 ± 0.2	35.1 ± 0.3	1806 ± 32
1	46.0 ± 0.1	29.2 ± 0.5		32.1 ± 0.1	1089 ± 7
2	46.2 ± 0.2	28.7 ± 0.1		31.4 ± 0.3	569 ± 11
3	46.3 ± 0.1	28.4 ± 0.3		31.3 ± 0.4	380 ± 10
4	46.3 ± 0.1	28.3 ± 0.1		30.8 ± 0.7	279 ± 54
5	46.1 ± 0.2	27.6 ± 0.3		31.5 ± 0.9	226 ± 18
10				30.0 ± 0	122 ± 30
15				30.2 ± 0.5	80 ± 11
20				29.1 ± 0.1	64 ± 11

Three determinations on each of three separate samples were carried out. The averages and standard deviations are reported

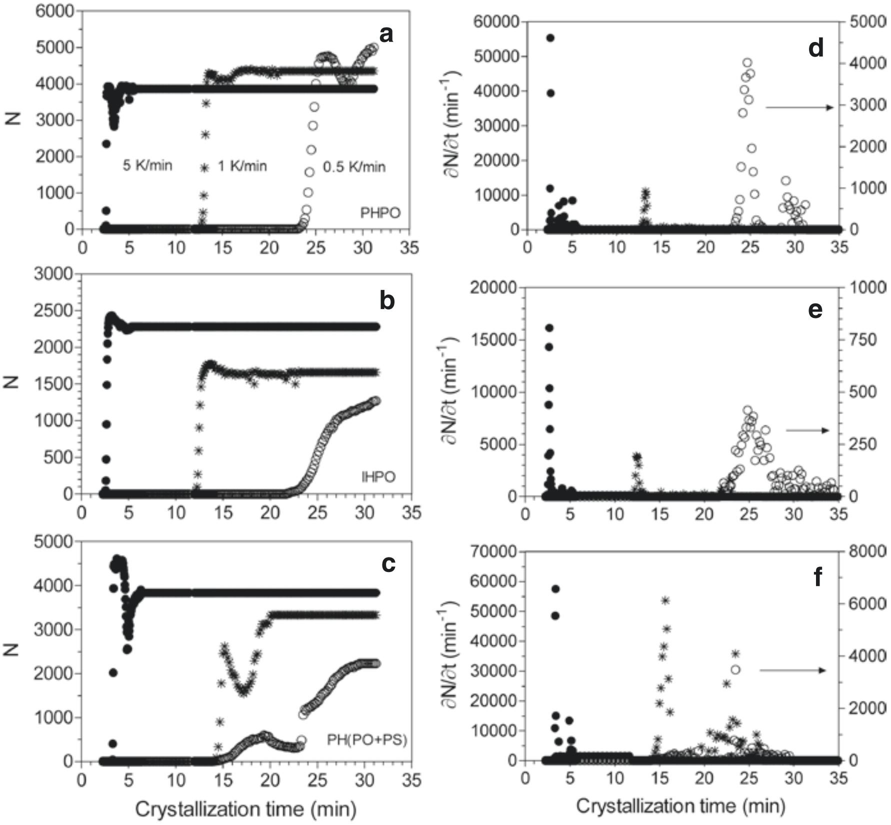
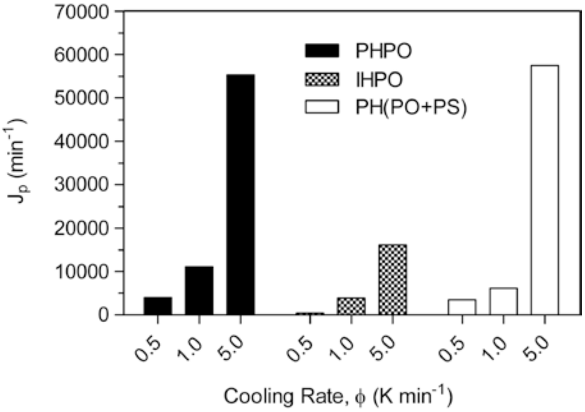


Fig. 9.5 Changes in number of crystals as a function of crystallization time for three multicomponent fat systems (**a**–**c**), and the corresponding first derivatives of the patterns (**d**–**f**)

Fig. 9.6 Effect of cooling rate on the peak nucleation rate of the fats determined from the maxima in Fig. 9.5 **b, d, e, f**



in the peak nucleation rate. For identical crystallization conditions, the nucleation rate as determined by polarized light microscopy was plotted against the nucleation rate as determined by taking the inverse of the induction time. If the two methods were exactly identical, the plot would show a diagonal line with a slope of 1 and a correlation coefficient of 1. The correlation coefficient (r^2) obtained was 0.82 and the slope 1.1 ± 0.2. Thus, there is reasonable agreement between the two methods, thus validating the approximation $J \sim \frac{1}{t_c}$.

As shown in Table 9.2, crystallization temperatures obtained for the same sample at the same cooling rate were higher for the PTA than for the DSC. Since crystallization is carried out by decreasing the temperature until the set crystallization temperature T_{set} is reached, this suggests that the PTA detects the onset of crystallization much earlier than the DSC. The logical conclusion is that the PTA technique is more sensitive than the DSC technique at detecting the appearance of the first crystals (onset of crystallization). Given the high sensitivity of the PTA, it is reasonable to assume that the induction times determined using this technique are reasonable estimates of the induction times of nucleation.

The induction times were shorter when the samples were crystallized at higher cooling rates. This observation applies to all 5 fat samples. This is an expected result since the higher the cooling rate, the shorter the time required to reach the crystallization temperature. Also, it can be observed from Table 9.2 that when the cooling rate is high, the temperature of the onset of crystallization T_c is lower than with a low cooling rate. This validates the assumption that it is the *time of exposure* to supercooling that is of key importance.

Curve fitting of the model to the data was carried out using equations in the following form:

$$J = \left(J_{\max} - J_{\lim}\right)e^{-k\beta} + J_{\lim} \tag{9.16}$$

where $J_{\lim}$ is the limiting nucleation rate as $\beta \to \infty$. We found this was necessary with nucleation rate data obtained from the inverse of an induction time. The introduction

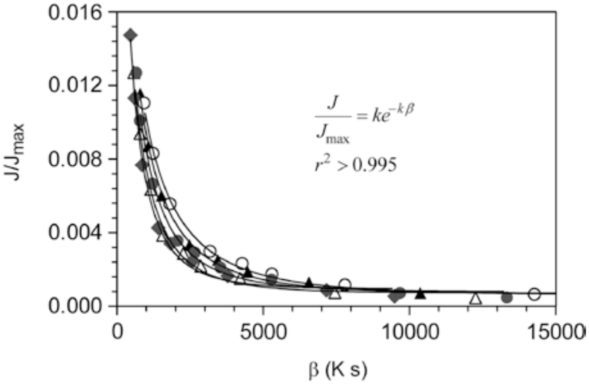


Fig. 9.7 Variation of the normalized nucleation rate (J/J_{max}) as a function of. the supercooling-time exposure, namely the effective supercooling (β) for the different systems studied

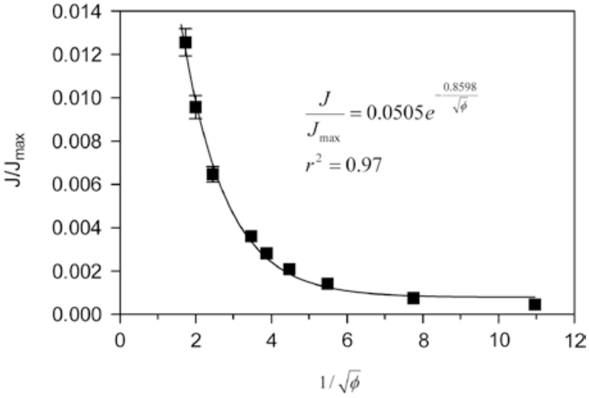


Fig. 9.8 Variation of the normalized nucleation rate as a function of the cooling rate for all systems grouped together

of the limiting nucleation rate (as opposed to a decay to zero nucleation rate) lead to a statistically significant improvement in the fit of the model to the data ($P < 0.05$).

The fit of the data to the model was excellent, as can be appreciated for the different systems in Fig. 9.7. It is interesting to notice that when the normalized nucleation rate was plotted as a function of $\frac{1}{\sqrt{\phi}}$, (Fig. 9.8), all lines collapsed onto a single master curve and were not significantly different from each other ($P > 0.05$).

It is possible to determine the energy of activation for nucleation using the approach previously developed in this section. Briefly, the molar activation energy of nucleation can be calculated as explained before as $Q = Z * X * MW$, which in this case was determined to be $X = 0.8598\ \mathrm{K}^{1/2}\mathrm{s}^{1/2}$. The average molecular weight MW of the fat was assumed to be 800 g/mol. Table 9.3 shows the values for k, Z, $J_{\max}$ and Q_M determined for each sample. Although the $J_{\max}$ of the samples were not significantly different from each other, a trend could be observed between the activation energy Q_M and the maximum nucleation rate $J_{\max}$. The lower the Q_M value, the higher the nucleation rate, as expected from kinetic theory, if Q_M represents the energy of activation for the nucleation process.

Table 9.3 Nucleation kinetic parameters (k, Z), energy of activation per unit mass (Q_m) and per mol (Q_M), and maximum nucleation rate (J_{max}) for the systems studied

System	k^1 ($K^{-1/2}s^{-1/2}$)	Z ($J\ g^{-1}\ K^{-1/2}s^{1/2}$)	Qm (J/g)	QM (kJ/mol)	Jmax^1 (s^{-1})
AMF	0.0775^{ab} (0.00480)	36.5	31.4	25.1	1.32^{a} (0.079)
PHPO	0.0622^{bc} (0.00339)	45.5	39.1	31.3	1.23^{a} (0.056)
PO	0.0558^{c} (0.00329)	50.7	43.6	34.9	1.20^{a} (0.056)
IHPO	0.0848^{a} (0.00510)	33.4	28.7	23.0	1.26^{a} (0.070)
PH(PO/PS)	0.0701^{ac} (0.00557)	40.3	34.6	27.8	1.23^{a} (0.084)
Average ($n = 5$)	0.070 (0.0050)	41.3	35.5	28.4	1.25 (0.020)

[1] Values reported are the average and standard error ($n = 9$). Values with the same superscript letter within a column are not significantly different from each other ($P > 0.05$)

One interesting finding is that the $J_{\max}$ values were not statistically different from each other ($P > 0.05$), with an average value of 1.25. The significance of this is that there is evidence that 1.25 can be used as a “universal” $J_{\max}$ constant for the five fat materials described. As such, it is thus possible to predict the nucleation rate, or induction time of nucleation, of triacylglycerols under nonisothermal conditions if the energy of activation is known:

$$J = \frac{1}{t_c} = 1.25ke^{-\frac{Q_m}{Z\sqrt{\phi}}} \tag{9.17}$$

The parameters required are listed in Table 9.3; the energy of activation in the above expression should be used with units of kJ/mol.

Moreover, a reasonable approximation of the nucleation rate of fats as a function of cooling rate can be obtained using:

$$J\left[s^{-1}\right] = \frac{1}{t_c} = 0.0632e^{-\frac{0.86}{\sqrt{\phi}}} \tag{9.18}$$

It is worth pointing out that the entire kinetic characterization of nucleation was carried out using simple light scattering device. The PTA analyzer was used to characterize the melting as well as the crystallization behavior. This procedure could thus be completely automated and a standard method developed for the characterization of the non-isothermal nucleation behavior of fats.

The characterization of nucleation kinetics under non-isothermal conditions was achieved by parametrization of the data considering both time and supercooling effects. A time-dependent supercooling parameter was defined and found to be related to the nucleation rate in a simple exponential decay fashion. The parametrization procedure used in the analysis of the nucleation kinetics of the fats crystallized at different cooling rates allowed for the determination of a cooling-rate independent energy of activation for nucleation.

Bibliography

1. Kashchiev D (2000) Nucleation at variable supersaturation. In: Kashchiev D (ed) Nucleation: basic theory with applications. Butterworth Heinemann, Oxford, pp 279–289
2. Marangoni AG, Wesdorp LH (2013) Structure and properties of fat crystal networks, 2nd edn. CRC Press, Boca Raton
3. Marangoni AG, Tang D, Singh AP (2006) Nonisothermal nucleation of triacylglycerol melts. Chem Phys Lett 419:259–264
4. Marangoni AG, Aurand T, Martini S, Ollivon M (2006) A probabilistic approach to model the non-isothermal nucleation behavior of triacylglycerol melts. Cryst Growth Des 6:1199–1205
5. Rogers MA, Marangoni AG (2008) Non-isothermal nucleation and crystallization of 12-hydroxystearic in vegetable oils. Cryst Growth Des 8:4596–4601
6. Rybin EN (2000) On the kinetics of non-isothermal nucleation. Colloid J 65:230–236

Chapter 10
Implementation of the van Smoluchowski Model for Protein Aggregation Kinetics: Cold-Gelation of Heated Whey Protein Isolate

10.1 Introduction

Paul Flory defined a substance as a gel "if it has a continuous structure with macroscopic dimensions that is permanent on the time scale of an analytical experiment and is solid-like in its rheological properties." Gels form when particles (in the case of particulate gels) or polymers (in the case of polymer gels) form an interconnected network with elastic properties. This network has a high surface area capable of entrapping large volumes of liquid. In many instances, solid-like gels with 99% liquid can be made.

Gels are responsible for many aspects of food microstructure and rheological properties in proteinaceous foods such as eggs, dairy (cheese) and soy (tofu) products. For this reason, characterizing the mechanisms responsible for their assembly into such networks has important implications for food quality. By determining the mechanisms and kinetics of aggregation, and quantitatively describing the geometry of the resulting network, food technologists will be better able to control macroscopic functionality of food products, such as texture and perceived color.

The "particle" in a particulate gel is a colloid-sized supramolecular aggregate formed from the aggregation of proteins. The interaction between these colloidal particles results in the formation of a continuous 3-dimensional network capable of entrapping liquid. The process of aggregation determines many of the gel's properties. Colloidal particles interact with each other in a stochastic, random, fashion, forming a diverse range of microstructures, from completely random to fractal. The network formed entraps liquid in the inter-particle and inter-cluster spaces. Alternatively, protein particles can also aggregate linearly, thus forming strand-like structures. These supramolecular strands can then interact with each other, gel, and entrap the solvent within.

Proteins must often undergo processing before they can be used for gelling applications. For example, in cheese-making, rennet (proteolytic enzymes) is required to cleave off κ-casein from the major milk protein casein. κ-casein is charged and

A.G. Marangoni, *Kinetic Analysis of Food Systems*,
DOI 10.1007/978-3-319-51292-1_10

glycosylated and is present as a “hairy layer” on the exterior of the casein micelle. κ-casein stabilizes the colloidal system via electrostatic repulsion and steric hindrance. Once cleaved, “bald” casein micelles can interact with other particles via hydrophobic interactions and form a gel.

Gelation can also be induced by heating. Heating causes the unfolding of protein molecules. Denaturation, in turn, exposes specific functional groups. These groups can interact with each other, which in turns leads to the formation of crosslinks between polymer chains, eventually resulting in the formation of a network.

Gelation can also be induced by covalent crosslinking. An example of this process involves the enzyme transglutaminase, which catalyzes the formation of amide bonds between carboxylic acid and amine side groups. This crosslinked gel is an example of a “chemical” gel, rather than the “physical” gels discussed above.

Modelling the gelation process is not simple and requires both excellent experimental techniques and the appropriate theoretical models. Here we briefly present two such models.

10.2 van Smoluchowski’s Theory of Rapid Coagulation (Physical Gels)

van Smoluchowski considered the rapid coagulation of spherical colloidal particles to be proportional to the number of collisions of single particles with each other or with existing higher aggregates. Collisions between these particles are assumed to be controlled only by the Brownian motion of the interacting particles, i.e., diffusion. The model also defines a sphere of interaction for the aggregating particles, with the sphere of interaction having a radius twice that of the interacting particles. If two particles approach to a distance smaller than their radii of interaction, i.e., if they “touch” as far as van der Waals forces are concerned, they would irreversibly aggregate. The van Smoluchowski model only applies to spherical particles. It models aggregation reasonably well at the initial stages of protein aggregation when the system is still monodisperse (particles all have a similar size).

van Smoluchowski described colloidal aggregation as a second order irreversible process of the type:

$$[i]+[j]\xrightarrow{k_S}[i+j] \tag{10.1}$$

where [i] and [j] correspond to the monomer and/or any aggregate reacting with each other, and k_S is the Smoluchowski rate constant for this reaction.

For example, in the case of tetramers being formed, dimers can react with dimers, and trimers with monomers to yield a tetrameric reaction product. Moreover, dimers are formed from monomer-monomer interactions, and trimers from dimer-monomer interactions, namely:

$$M + M \rightarrow M_2$$

$$M + M_2 \rightarrow M_3$$

$$M_2 + M_2 \rightarrow M_4$$

$$M + M_3 \rightarrow M_4$$

In general, the simplified differential equation which describes changes in the concentration of the k^{th} aggregated species (C_k) in time during this random aggregation process, where all rate constants are equal, has the form:

$$\frac{dC_k}{dt} = K\left[\frac{1}{2}\sum_{j=1}^{k-1} C_{k-j}C_j - C_k\sum_{j=1}^{\infty} C_j\right] \tag{10.2}$$

where k_S represent is the Smoluchowski rate constant for the reaction between an "*i-mer*" and a "*j-mer*" to form a "*k-mer*" and viceversa. If the initial concentration of monomers is C_0, conservation of particles requires $\sum_{k=1}^{\infty} C_k k = C_0$.

If only monomers are present at $t = 0$, Smoluchowski was able to provide an analytical solution to Eq. 10.2, with the form

$$C_k\left(t\right) = \frac{C_0\left(k_S C_0 t/2\right)^{k-1}}{\left(1 + k_S C_0 t/2\right)^{k+1}} \tag{10.3}$$

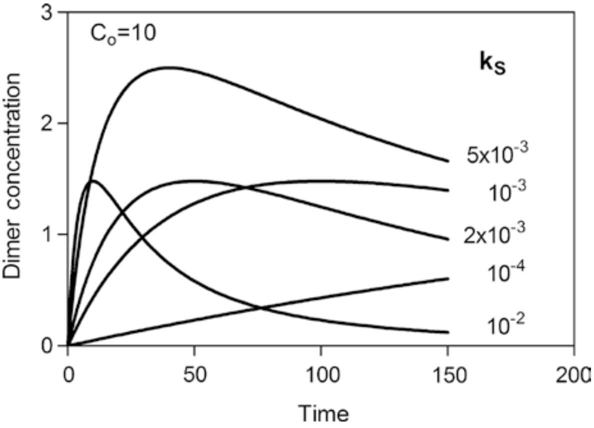
This equation describes the dynamic where the concentration of monomer continuously decreases, while the concentration of all *k-mers* increases and then decreases in time. For example, if only the monomer to dimer transition is considered, changes in the concentration of dimer as a function of time is given by

$$C_2 = \frac{C_0\left(k_S C_0 t/2\right)}{\left(1 + k_S C_0 t/2\right)^3} \tag{10.4}$$

A simulation of the curves obtained for a putative monomer to dimer transition for different values of the Smoluchowski rate constant are shown in Fig. 10.1. Notice the characteristic initial increase in dimer concentration, which is eventually followed by a decrease as the concentration of trimers takes over.

Experimentally, it is very difficult to determine the concentration of all these aggregates in time. However, one can focus on the early stages of the reaction, where only monomer to dimer formation is taking place and/or monitor decreases in monomer concentration, which would be modelled by

Fig. 10.1 Simulations of the Smoluchowski model of colloidal aggregation at different rate constants



$$C_1 = \frac{C_0}{\left(1 + k_S C_0 t / 2\right)^2} \tag{10.5}$$

Thus, fits of experimentally determined increases in dimer concentration or decreases in monomer concentration would yield an estimate of the Smoluchowski rate constant for the gelation process (see example).

This rate constant is proportional to the relative diffusion coefficient of two spherical single particles, $D_{AA} = 2D_A$, for equally sized particles, or $D_{AB} = D_A + D_B$, for unequally sized particles. The rate constant is also proportional to the radius of the sphere of interaction, $R_{AA} = 2a$, for equally sized spheres, and $\mathrm{R}_{AB} = \frac{\left(a+b\right)^2}{ab}$, for unequally sized spheres, where *a* and *b* are the corresponding particle radii. For unequally sized spheres, the Smoluchowski rate constant has the form:

$$k_s = 2\pi R_{AB} D_{AB} \tag{10.6}$$

while for equally sized spheres, this equation has the form:

$$k_s = 2\pi R_{AA} D_{AA} \tag{10.7}$$

with units of $[M^{-1}\ s^{-1}]$. It is therefore possible to derive useful parameters from knowledge of the Smoluchowski rate constant of aggregation.

If the particles are assumed to be spherical, the Stokes-Einstein relationship can be used to show the dependence of the Smoluchowski rate constant on the viscosity of the medium. The diffusion constant is related to the viscosity of the medium:

$$D_A = \frac{k_B T}{6\pi\eta a} \tag{10.8}$$

where k_B is Boltzman's constant, T is the absolute temperature, η is the viscosity of the medium and a is the radius of particle A.

For two equally sized aggregating particles, A, the effective diffusion constant D_{AA} is given by:

$$D_{AA} = \frac{k_B T}{6\pi\eta}\frac{1}{a} + \frac{k_B T}{6\pi\eta}\frac{1}{a} = \frac{2k_B T}{6\pi\eta a} \tag{10.9}$$

For two differently sized aggregating particles, A and B, the diffusion constant D_{AB} is given by:

$$D_{AB} = \frac{k_B T}{6\pi\eta}\frac{1}{a} + \frac{k_B T}{6\pi\eta}\frac{1}{b} = \frac{k_B T}{6\pi\eta}\frac{a+b}{ab} \tag{10.10}$$

Since $R_{AB}=a+b$, substituting Eqs. 10.8 into 10.6 would yield the following expression for the Smoluchowski rate constant:

$$k_s = \frac{1}{3}\frac{k_B T}{\eta}\frac{(a+b)^2}{ab} \tag{10.11}$$

If the particle radii were equal, $a=b$, then:

$$k_s = \frac{4}{3}\frac{k_B T}{\eta} = 4\pi D_{AA} a \tag{10.12}$$

Equations 10.11 and 10.12 indicate that the greater the diffusivity and radii of interaction of the interacting particles, the greater the rate constant of the process would be. This would translate into a greater probability for the aggregation event to occur. A high viscosity, however, would impede the diffusion of reacting particles through the medium. A low-medium viscosity, however, would enhance aggregation. The rate constant for the process is greater for a polydisperse system, since the radii of interaction of the aggregating particles are larger.

As stated before, the van Smoluchowski approach works well for initially monodisperse or low initial polydispersity systems. An important modification is that of corrections for slow aggregation. Previously we assumed that every collision was effective and led to an aggregation event. This is not the case in reality. For example, there may be an electrostatic energy barrier that has to be surmounted before coagulation takes place. To compensate for this effect, a probability factor for a "productive collision" is introduced into the scheme. The Smoluchowski rate constant for two equal interacting spherical particles then takes the form:

$$k_s = \alpha 2\pi R_{AA} D_{AA} \tag{10.13}$$

The factor α is the "collision probability" or fraction of collisions that cause aggregation. This factor eventually becomes a "fitting parameter" and floated during fits. If this factor is to be introduced, the user should have an independent way of obtaining it, i.e., not through the fit to the data, and should be constrained to be a constant.

10.3 Polyfunctional Condensation Model (Chemical Gels)

Rather than considering aggregation as a random, solely diffusion-limited process, some researchers have attempted to describe certain protein aggregation processes as an actual chemical polymerization reaction involving monomeric units. The rate limiting step of the reaction is not the diffusion of the colloids, but rather the chemical reaction that takes place between colloidal particles. An example of this process is the Ca^{+2} mediated aggregation of caseins or the transglutaminase-catalyzed crosslinking of proteins.

In this approach, the concept of functionalities, or reactive groups, is borrowed from small molecule polymerization theory. This concept of a functional group is rather nebulous for a macromolecule such as a protein, which may contain several such functional groups in variable quantities. At least two functionalities are required for the colloidal particle to form a gel network consisting of branched aggregates. As well, a protein in its native conformation will not have its functional groups exposed, thus these are unavailable for reaction. It is often necessary to first denature a protein in order to expose its functional groups. This approach allows for the mechanistic description of protein aggregation processes that is similar to polymerization reactions.

The polyfunctional model is similar to the van Smoluchowski approach, in that the aggregation process is considered a second order process:

$$A + A \rightarrow A_2$$

In this theory, A stands for the concentration of functionalities instead of the concentration of reactant molecules. The change in the concentration of functional groups available to undergo reaction as a function of time (C_t) has the following form:

$$C_t = \frac{C_0}{1 + kC_0 t} \tag{10.14}$$

where C_0 is the initial concentration of functionalities and k is the second order rate constant of the process in units of $M^{-1}t^{-1}$.

Gordon [2] derived an expression relating the weight average molecular weight (M_w) and the number average of a polymerizing material (M_n) to the fraction of functionalities that have reacted (α):

$$M_w = M_0 \left(\frac{1 + \alpha}{1 - (f - 1)\alpha} \right) \tag{10.15}$$

$$M_n = M_0\left(\frac{2}{2 - f\alpha}\right) \tag{10.16}$$

where M_o is the molecular weight of the monomer and f is the number of functionalities per monomer.

Since α is the proportion of reacted functionalities, then $\alpha = \left(1 - \frac{C_t}{C_0}\right)$. Substituting this into Eq. 10.15 yields the expression:

$$M_w = M_0\left(\frac{1 + 2kC_0t}{1 - kC_0\left(f - 2\right)t}\right) \tag{10.16}$$

Hence, by monitoring changes in the molecular weight of a protein during aggregation, it is possible to derive a reaction constant for aggregation and the number of functionalities present per protein monomer. Changes in the molecular weight of an aggregating protein as a function of time can be fitted to the function above by nonlinear regression, and the parameters derived. On the other hand, it is possible to assume the value of f to be a constant in order to unambiguously determine the rate constant. Fixing the value of f as a constant in the fitting procedures would be necessary for the unambiguous determination of the rate constant.

10.4 Aggregation of Whey Protein Isolate

Whey protein isolate (WPI) is a protein fraction derived from whey, a by-product of the cheese manufacturing process. It contains approximately 90% protein. Whey protein isolates are frequently used as gelling agents in functional food applications. Used in this manner, WPIs are added to a food formulation at a high concentration. The suspension is then heated to high temperatures to induce gelation. The main drawbacks of this method are the undesirable physical and chemical changes that occur during heating. As well, this method typically forms opaque gels which are undesirable. The gel is also less elastic and the material can be described as "crumbly". A new process, cold gelation, is currently being investigated as a suitable replacement for the conventional gelation process. Cold gelation is carried out at low temperatures, thus, undesirable physical and chemical changes are avoided. Cold gelation results in more transparent and elastic gels than those formed by conventional gelation. The process of cold gelation involves first making a low-ionic strength sol from whey proteins in water. If the protein concentration is less than 12%, heating the suspension will not result in aggregation. The sol is cooled and then added to the food formulation.

Gelation is induced by the addition of salts such as Ca^{2+} or Na^+. The effect of calcium on gelation is two-fold. As a positively charged ion, Ca^{2+} can neutralize or disperse negative charges, reducing electrostatic repulsion between protein molecules. This promotes more random aggregation of whey proteins, resulting in more opaque gels. However, gels formed with calcium are often observed to be transparent. This indicates that salting-out is not the sole effect of calcium on gelation. Calcium ions, as divalent ions, can cross-link negatively charged residues on proteins. This cross-linking effect slows down the rate of aggregation, resulting in a gel with a more ordered network structure. The balance between these two processes determines the final properties of the gel.

The rate of the processes involved in protein coagulation is important in determining the final properties of the product. In particular, the rate at which the protein is denatured relative to its rate of aggregation is important for determining, among other things, the opacity of the resulting network. If the rate of denaturation is relatively rapid compared to the rate of aggregation, a continuous and translucent hydrogel is formed. On the other hand, if the rate of aggregation is more rapid than the rate of denaturation, then opaque particles with less water-holding capacity are formed. The rates of these processes are, in turn, affected by properties such as protein concentration, pH, ionic strength and heating conditions. In brief, opaque gels are formed from random aggregation, which is the result when the rate of aggregation is high. In contrast, transparent gels are formed from linear aggregation, which occurs when the rate of aggregation is low.

The effects of pH are most pronounced when the protein is close to its isoelectric point. At the isoelectric point, the charged moieties of the protein are uncharged due to protonation or deprotonation effects. At zero charge, hydrophobic and van der waals interactions predominate and the protein aggregates rapidly due to a lack of electrostatic repulsion. This rapid aggregation promotes random aggregation. At the pHs away from the isoelectric point, aggregation is slow as high electrostatic repulsion will hinder the aggregation of proteins. Aggregation is linear under such conditions. In the case of calcium-induced WPI cold gelation at pH 7, the following factors have been shown to induce the formation of more transparent gels: high pre-heating temperature, long pre-heating time, high protein concentration, low gelation temperature and low calcium levels. Increasing calcium concentration at pH 7 results in increasing gel opacity.

10.5 Aggregation Kinetics Example: The van Smoluchowski Model Implemented

In the following section, the kinetics of the cold-gelation of whey protein isolate in the presence of calcium are quantified using the van Smoluchowski model, and the rate constant dependence on calcium concentration determined. A 4% whey protein isolate solution at pH 7 was made and then heated at 80 °C for 30 min. The solution was then cooled and stored overnight at 3 °C. Cold-gelation was induced when this protein solution was mixed with different $CaCl_2$ solutions resulting in 2% WPI solutions with differing final calcium concentrations. The gelation time was determined as the time at which no flow of sample was observed when sample-containing tubes

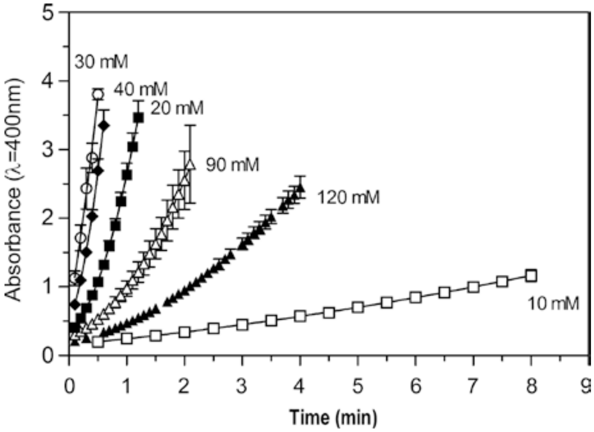


Fig. 10.2 Increase in the absorbance at 400 nm (A_{400}) as a function of time of a 2% (w/v) heated whey protein isolate solution gelled at 3 °C in the presence of different calcium chloride concentrations

were inverted. The kinetics of the aggregation process were quantified by measuring the absorbance of the solutions at 400 nm as a function of time. The increase in absorbance of the sample at this wavelength was mainly due increases in the turbidity of the samples upon protein aggregation.

Figure 10.2 displays the change in absorbance at 400 nm (A_{400}) for the different samples as a function of time. As can be observed from the progress curves, the rate of aggregation is greater at higher calcium concentrations. The rate at 90 mM of calcium was greater than the rate at 10 mM of calcium. However, it is interesting to note that the rate at 40 mM is greater than the rate at 90 mM, suggesting that too high of a calcium concentration may in fact slow down the aggregation process.

The reader should note at this point that we would have a problem if one tried to fit the absorbance data instead of concentration to models. Firstly, the sample absorbs light due to a turbidity effect rather than by a specific chromophore. This absorbance is due to light scattering by particles in suspension and thus influenced by more factors than just solution concentration. Thus, the relationship of this absorbance to colloid concentration is not clear. Thus, it is necessary to explore the possibility of transforming this absorbance data to scattered intensity, which then could be used to obtain an approximation to colloid concentration.

In a typical static light scattering experiment, the amount of scattered light (I_S) in a suspension is given by:

$$I_S = I_0 K C M_w P(\theta) \frac{V_S}{d} \tag{10.17}$$

Where I_0 is the intensity of the incident light beam, K is a constant which depends on the experimental setup and sample (wavelength of the light used, refractive index of the sample and solvent), C is the weight concentration of the sample (kg/L), M_w is the molecular weight of the scattering particle, V_S is the scattering volume, d is the distance from the scattering particles to the detector and $P(\theta)$ is the form factor dependent on the size and shape of the particles. The form factor at $\theta = 0°$ is equal to unity ($P(\theta = 0°) = 1$). For the purpose of the discussion that follows, I will express the relative scattered intensity (I_S/I_0) at $\theta = 0°$ for a sample with a scattering particle of constant molecular weight as:

$$\frac{I_S}{I_0} = FC \tag{10.18}$$

where $F = KM_w P(\theta)\frac{V_S}{d}$.

Now, considering and assuming that the absorbance of the samples increased due to turbidity increases upon formation of dimers only, we can determine the relative amount of scattered light by:

$$\frac{I_S}{I_0} = \frac{I_0 - I_{absorbed}}{I_0} \tag{10.19}$$

This is quite convenient since $A = -\log_{10}\frac{I_{absorbed}}{I_0}$, which means that the relative amount of scattered light can be determined from absorbance measurements as $\frac{I_S}{I_0} = 10^{-Absorbance}$. Armed with this, we could then obtain an estimate of the concentration in solution using Eq. 10.18. However, the reader should be warned since this concentration is an "apparent" parameter unless we determine the value of the constant F in Eq. 10.18, and ensure that our assumptions are correct, such as the constancy in molecular weight. However, this does not preclude the *careful* use of this approximation to colloidal concentration.

On that note, the fit of the van Smoluchowski model to this transformed data was remarkable (Fig. 10.3), with correlation coefficients above 0.999. This partially demonstrates that the approach has merit and that the theory can explain the observed behaviour quite closely. The Smoluchowski rate constants (k_S) derived from fits of the monomer-to-dimer model to the data were plotted as a function of calcium concentration (Fig. 10.4). A maximum in the aggregation rate was observed at a calcium concentration of 30 mM. This pattern agreed very well with the independently-determined experimental gelation time, which showed a minimum at 30 mM calcium. As a technical point related to data-fitting, the transformed data had to be adjusted before proper model fitting could be carried out. Since the progress curves did not start at $x = 0$ or $y = 0$, some adjustment was necessary so that the first data point was positioned at $x = 0$ and $y = 0$. Time data was adjusted at $t_{zeroed} = t_i - t_1$, where the time value of the first data point (t_1) was subtracted from all others (t_i). These new time values were labelled "zeroed time" (Fig. 10.3). The y-values were also adjusted by baseline correction, where the first (and lowest) transformed absorbance value was subtracted from all other values. The new transformed absorbance values were labelled as "baseline corrected" (Fig. 10.3).

The maximum aggregation rate observed at a calcium concentration of 30 mM, or ionic strength, μ, of 75 mM could arise due to the fact that at low calcium

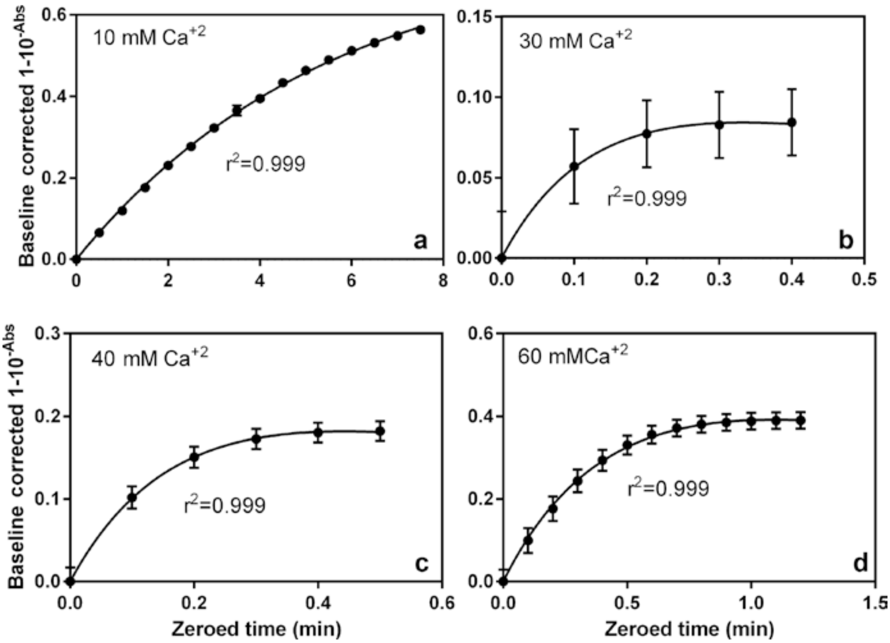


Fig. 10.3 Baseline-corrected (average of first *y*-value was substracted from all other absorbance values) and zeroed (time of first measurement was substracted from all time values) apparent relative intensities of scattered light (I_s/I_0) as a function of time data during gelation of heated whey protein isolate solution gelled at 3 °C in the presence of different calcium chloride concentrations. Symbols represent averages of 5 replicates and their corresponding standard errors. Solid lines represent the fit of the von Smoluchowski model to the data considering only a monomer to dimer transition

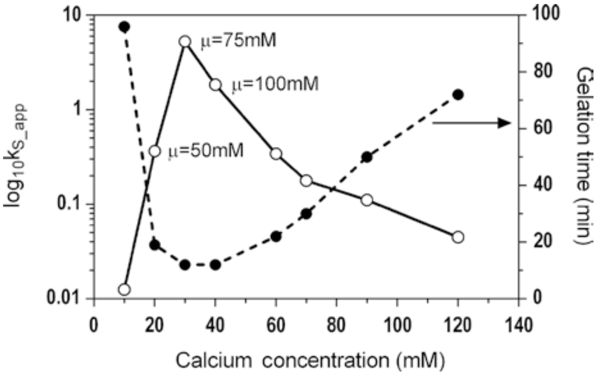


Fig. 10.4 Changes in the apparent von Smoluchowski rate constant (k_s) as a function of calcium concentration for heated whey protein isolate solution gelled at 3 °C. The gelation time (zero flow when tube inverted) was also plotted here to demonstrate that both k_s and gelation time follow similar trends

concentrations electrostatic repulsion prevents protein colloids from approaching sufficiently for "sticking", and aggregation to take place. Possibly under these low ionic strength conditions, charge dispersion is weak (larger Debye lengths), and electrostatic repulsion prevents proteins from aggregating, thus resulting in a lower rate of aggregation. The Debye length (κ^{-1}) is the measure of a charge carrier's net electrostatic effect in solution, and how far those electrostatic effects persist. A longer Debye length means that electrostatic charges exert their effect over longer distances. The Debye length, or radius, decreases with increasing ionic strength, $\kappa^{-1} \sim \mu^{-1/2}$. The slow kinetics at high calcium concentrations could be due to ionic shielding and charge dispersion of any reactive negatively charged group on the protein by excess calcium ions. In this case, a further decrease in Debye length beyond the 30 mM calcium maximum, would prevent the calcium ionic bridge crosslinking reaction between negatively charged groups from taking place. The negatively charged groups and positively charged calcium would never "feel" each other.

In this chapter we have implemented van Smoluchowski's theory in the study of protein aggregation kinetics and shown its utility for the case of cold-gelation of heated whey protein isolate. Possibly a specific rate constant is associated with the most desirable structure. Further studies could include a study, based on a prediction, of the effects of solvent viscosity on the Smoluchowski rate constant (see model). Addition of a polysaccharide to the solution would affect the aggregation process, which in turn would affect the mechanical and optical properties of the gel. Correlating k_S to Debye lengths would be another fruitful research avenue. The added benefit of these investigations is that they would be hypothesis-driven and based on a mathematical prediction of system behaviour, which in my view, is always the best way to approach scientific problems.

Bibliography

1. Goodisman J, Chaiken J (2006) Scaling and the Smoluchowski equations. J Chem Phys 125:074304–074307
2. Gordon M (1962) Good's theory of cascade processes applied to the statistics of polymer distributions. Proc Roy Soc A 268:340–256
3. Holthoff H, Smitt A, Fernandez-Barbero A, Borkovec M, Caberizo-Vilchez MA, Schurtenberger P, Hidalgo-Alvarez R (1997) Measurement of absolute coagulation rate constants for colloidal particles: comparison of single and multiparticle light scattering techniques. J Colloid Interface Sci 192:469–470
4. Hongsprabhas P, Barbut S, Marangoni AG (1999) The structure of cold-set whey protein isolate gels prepared with Ca^{+2}. Lebensmittelwissenchaft und Technologie 32:196–202
5. Jullien R (1992) The application of fractals to colloidal aggregation. Croat Chem Acta 65(2):215–235
6. Marangoni AG, Barbut S, McGauley SE, Marcone M, Narine SS (2000) On the structure of particulate gels – the case of salt-induced cold gelation of heat-denatured whey protein isolate. Food Hydrocoll 14:61–74

7. Ogendal L (2016) Light Scattering: a brief introduction. University of Copenhagen. http://igm.fys.ku.dk/~lho/personal/lho/LS_brief_intro.pdf. Accessed 10 Oct 2016.
8. Parker TG, Dalgleish DG (1977) The potential application of the theory of branching processes to the association of milk protein. J Dairy Res 44:79–84
9. van Smoluchowski M (1916) Drei Vorträge über Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen Zeitschrift für Phys 17:557–571
10. van Smoluchowski M (1917) Versuch einer mathematischen theorie der koagulation kinetic kolloider losungen. Zeitschrift für Phys Chemie 92:129–168