

UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Ciencias Físicas
Departamento de Óptica



TESIS DOCTORAL

Models for chiral amplification in spontaneous mirror symmetry breaking.
Modelos de amplificación quiral en ruptura espontánea de simetría especular

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

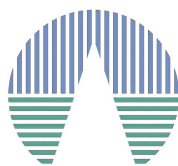
PRESENTADA POR

Celia Blanco de Torres

Director

David Hochberg

Madrid, 2014



CENTRO DE ASTROBIOLOGÍA
ASOCIADO AL NASA ASTROBIOLOGY INSTITUTE



GOBIERNO
DE ESPAÑA



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



Instituto Nacional de
Técnica Aeroespacial

Models for Chiral Amplification in Spontaneous Mirror Symmetry Breaking

**Modelos de Amplificación Quiral
en Ruptura Espontánea de Simetría Especular**

Memoria que para optar al grado de Doctora en Física presenta

Celia Blanco de Torres

dirigida por: Dr. David Hochberg

Madrid, Marzo de 2014

Departamento de Óptica
Facultad de Ciencias Físicas
Universidad Complutense de Madrid

A mi yayo

"Science is the belief in the ignorance of experts"
R. P. Feynman

The work presented in this doctoral thesis has been funded through a Calvo-Rodés scholarship from the Instituto Nacional de Técnica Aeroespacial (INTA) and the Research project AYA2009-13920-C02-01: "Experimental and theoretical models for the abiotic emergence of chirality and its detection as the signature for evolutive systems in extraterrestrial materials", from the Ministerio de Ciencia e Innovación (Spain).

El trabajo presentado en esta tesis doctoral se ha financiado gracias a la concesión de una beca de Formación de Personal Investigador Calvo Rodés del Instituto Nacional de Técnica Aeroespacial (INTA) y al proyecto de investigación AYA2009-13920-C02-01: "Modelos experimentales y teóricos sobre la emergencia abiótica de la quiralidad y su detección como la firma de sistemas evolutivos en materiales extraterrestres", del Ministerio de Ciencia e Innovación.

Cover: Mirror images (enantiomers) of a chiral amino acid.

Portada: Imágenes especulares (enantiómeros) de un aminoácido quiral.

Agradecimientos

Hay tanta gente sin la que este trabajo no hubiese salido adelante, que seguramente me dejaré nombres. Sin duda, el primero al que tengo mucho que agradecer es a mi director de tesis, David Hochberg, por haberme dado la oportunidad de embarcarme en este proyecto, y por toda su dedicación y apoyo.

También tengo mucho que agradecer a los que han sido mis compañeros durante estos cuatro años en el Centro de Astrobiología: a Alicia, Viki y Laura, por haber estado siempre disponibles para ayudarme en todo lo que hiciese falta, a Pablo, Martin y Alfonso por tantas y tantas conversaciones dentro y fuera del CAB, a las dos Marías -María (Jaén) y María (Santiago)-, a Michael, Irene, Patxi, Jacobo, Vanessa, Enoma, Patri, Hector, Fernando... a todos ellos y a muchos más, por su ayuda, su consejo, su compañerismo, y lo mucho que he aprendido junto a ellos estos años. A los informáticos -Paco, Luis, Fernando y Jota- tengo que agradecerles mucha asistencia técnica, muchas sobremesas, y la impresión de algún que otro poster de manera express. A Luis y a Juan Ángel, enseñarme lo divertido que puede ser divulgar. Y a Esther y Tatiana, haber solucionado siempre mis despistes con paciencia y simpatía.

A los compañeros de la Universitat de Barcelona -Josep, Zoubir, Albert y Joaquim- tengo que agradecerles innumerables consejos en cada ocasión que hemos tenido de coincidir, además de su amabilidad en mi visita a Barcelona. Especialmente a Josep, he de agradecerle numerosas conversaciones en las que, con gran paciencia me ha explicado, enseñado y aconsejado sobre diversos temas. Colaborar con ellos ha sido una experiencia muy positiva y enriquecedora, de la que ha salido gran parte de esta tesis.

Por supuesto, nada de esto hubiese sido posible sin las personas que me rodean día a día, las que verdaderamente están ahí para celebrar conmigo los buenos momentos y para animarme en los no tan buenos. Por eso, estoy enormemente agradecida a mis amigos, por toda su ayuda incondicional y su constante empeño en mantener mi vida social activa. Y a mi familia, por el infito apoyo que me ha prestado durante todos estos años que he tardado en recorrer el largo camino para llegar hasta este punto, de manera muy especial a a mi madre, sin la que no sería como soy.

Celia Blanco de Torres
Madrid, Enero de 2014

Table of Contents

Resumen	3
Abstract	5
1 Introduction: chirality, biological homochirality and the origin of life	7
1.1 Introduction	7
1.2 Chance Theories	9
Amplification of tiny stochastic imbalances	9
1.3 Aim and outline of the thesis	11
2 Frank model	13
2.1 Introduction	13
2.2 Reversible Frank model	14
2.3 Results	15
Open system	17
Semiopen system	18
Closed system	18
2.4 Conclusions	18
3 Limited enantioselective (LES) autocatalytic model	21
3.1 Introduction	21
3.2 Results	23
LES cycle driven by an external reagent	23
LES in a temperature gradient	26
LES in a temperature gradient and compartmentalized autocatalysis	27
3.3 Conclusions	28
4 Chiral polymerization and copolymerization	31
4.1 Introduction	31
4.2 The polymerization model	32
4.3 The copolymerization model	34
4.4 Results	36
Interpretation of experimental data	40
4.5 Conclusions	43
5 Mirror symmetry breaking via β-sheet-controlled copolymerization	45
5.1 Introduction	45
5.2 Theoretical Methods I: Mass Balance	46
5.3 Theoretical Methods II: Probabilistic approach	48
5.4 Results	51
Additives of only one handedness	51
Racemic additives	53
5.5 Conclusions	54
Conclusions	57
Conclusiones	59

Resumen

Es un hecho empírico que hay un desequilibrio quiral absoluto (o ruptura de simetría especular) en todos los sistemas biológicos conocidos, donde los procesos cruciales para la vida, como la replicación, implican estructuras supramoleculares que comparten el mismo signo quiral (homoquiralidad). Estas estructuras quirales son proteínas, compuestas de amino ácidos encontrados como L-enantiómeros; y polímeros de ADN y ARN y azúcares, compuestos de R-monocarbohidratos. Basándonos en el hecho de que es químicamente imposible lograr una mezcla perfectamente racémica por motivos puramente estadísticos, podemos asumir la presencia de un pequeño exceso enantiomérico inevitable incluso en una mezcla racémica 'perfecta'. Entonces, en principio, el origen y la evolución de la homoquiralidad biológica y de la ruptura espontánea de simetría especular (Spontaneous Mirror Symmetry Breaking, SMSB) en los sistemas químicos y biológicos podrían explicarse teóricamente mediante un modelo en el cual un pequeño desequilibrio entre enantiómeros es amplificado, y el camino más probable para ello es mediante reacciones autocatalíticas. Nuestro principal propósito aquí es probar la capacidad de varios modelos autocatalíticos diferentes de amplificar un exceso enantiomérico inicial, ee_0 , incluso por debajo del desequilibrio esperado (es decir, $ee_0 < ee_{st}$).

En este contexto, introducimos y estudiamos el sistema autocatalítico más básico conocido conducente a la amplificación quiral - modelo de Frank - y su capacidad para amplificar las pequeñas desviaciones estadísticas iniciales respecto de la composición racémica ideal. Dependiendo las condiciones, esta amplificación puede ser sólo temporal, y puede ser interpretada como una excursión quiral en un espacio de fases dinámicas. Estas excursiones quirales se pueden estudiar a través de una combinación de análisis del espacio de fases, análisis de estabilidad y simulaciones numéricas, con el fin de determinar la forma en que dependen de si el sistema es abierto, semiabierto o cerrado. También la emergencia de homoquiralidad en autocatálisis enantioselectiva de compuestos que no pueden transformarse según el conjunto de reacciones de tipo Frank es discutida, con respecto al controvertido modelo de enantioselectividad (Limited Enantioselectivity, LES), integrado por autocatálisis enantioselectiva y no enantioselectiva acopladas, y que no puede conducir a SMSB en sistemas cerrados. Ya que la homoquiralidad biológica de los sistemas vivos implica grandes macromoléculas, la capacidad de amplificar (y transmitir a todo el sistema) esos pequeños excesos enantioméricos es estudiada para dos modelos cinéticos diferentes de polimerización y copolimerización quiral en sistemas cerrados al flujo de materia y de energía, y mostramos los resultados de ajustar el modelo de copolimerización a los datos experimentales de amplificación quiral de oligopéptidos.

Por último, tanto un modelo de equilibrio químico de copolimerización por control de plantillas como un enfoque probabilístico se presentan para describir los resultados de los escenarios experimentales de desimetrización inducida propuestos recientemente por Lahav y su grupo; estos mecanismos de amplificación quirales tienen lugar mediante polimerización controlada por láminas β racémicas, tanto en cristalitos en superficie como en solución. En este caso, la ruptura de simetría surge de la combinatoria, y no de fenómenos espontáneos (bifurcación). Estos efectos estadísticos /combinatorios/estocásticos no se deben a las pequeñas fluctuaciones inherentes quirales presentes en todos los sistemas químicos reales, sino a la oclusión aleatoria de aminoácidos anfitriones y huéspedes por los sitios quirales de la plantilla: los mecanismos propuestos aquí funcionan aún para las mezclas racémicas idealmente.

Abstract

It is an empirical fact that there is an absolute chiral imbalance (or mirror symmetry breaking) in all known biological systems, where processes crucial for life such as replication, imply chiral supramolecular structures, sharing the same chiral sign (homochirality). These chiral structures are proteins, composed of amino acids found as the left-handed enantiomers (L); and DNA, and RNA polymers and sugars, composed of right-handed (R) monosaccharides. Based on the fact that a perfect racemic mixture is chemically impossible to achieve on purely statistical grounds alone, we can assume the presence of an unavoidable tiny enantiomeric excess even in a 'perfect' racemic mixture. Then, in principle, the origin and evolution of biological homochirality and Spontaneous Mirror Symmetry Breaking (SMSB) in chemical and biological systems could be theoretically explained by a model in which a tiny imbalance of one enantiomer was amplified, and the most likely path for this is by autocatalytic reactions. Our main purpose here is to test the ability of some different autocatalytic models to amplify a tiny initial enantiomeric excess, ee_0 , even lower than the expected inherent imbalance (i.e., using $ee_0 < ee_{st}$).

Then, in this context, we introduce and study the most basic known autocatalytic system leading to chiral amplification- the Frank model - and its ability to amplify the initial small statistical deviations from the idealized racemic composition. Depending on the conditions, this amplification can be just temporary, and it can be interpreted as a chiral excursion in a dynamic phase space. These chiral excursions can be studied through a combination of phase space analysis, stability analysis and numerical simulations in order to determine how they depend on whether the system is open, semi-open or closed. Also the emergence of homochirality in enantioselective autocatalysis for compounds unable to transform according to the Frank-like reaction network is discussed with respect to the controversial limited enantioselectivity (LES) model composed of coupled enantioselective and non-enantioselective autocatalyses, which cannot lead to SMSB in closed systems. Since biological homochirality of living systems involves large macromolecules, the ability to amplify (and transmit to the entire system) those small initial enantiomeric excesses is studied for two different kinetic models of chiral polymerization and copolymerization in systems closed to matter and energy flow, and the results from fitting the copolymerization model to the experimental data on chiral amplification of oligopeptides are shown.

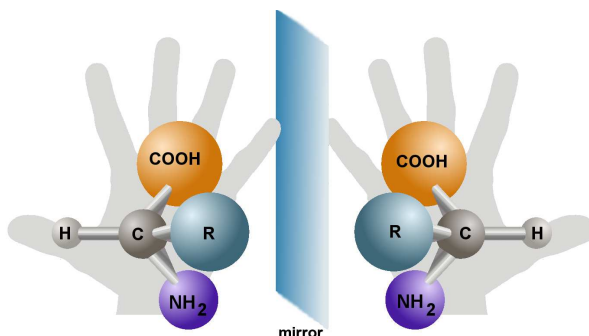
Finally, both a chemical equilibrium model of template-controlled copolymerization and a probabilistic approach are presented for describing the outcome of the experimental induced desymmetrization scenarios recently proposed by Lahav and co-workers; these chiral amplification mechanisms proceed through racemic β -sheet controlled polymerization operative in both surface crystallites as well as in solution. In this case, the symmetry breaking arises from combinatorics, not from spontaneous (bifurcation) phenomena. These stochastic/statistical/combinatorial effects are not due to the inherent tiny chiral fluctuations present in all real chemical systems, but are due rather to the random occlusion of host and guest amino acids by the chiral sites of the template: the mechanisms proposed here work even for ideally racemic mixtures.

Chapter 1

Introduction: chirality, biological homochirality and the origin of life

1.1 Introduction

A chiral molecule is one that is not superposable on its mirror image. Human hands are perhaps the most common example of chirality: The left hand is a non-superposable mirror image of the right hand; no matter how the two hands are oriented, it is impossible for all the major features of both hands to coincide (not even by pure rotations or translations). The term chirality is derived from the Greek word for hand ($\chi\epsilon\iota\rho$). Chirality plays an



important role in chemistry, where it usually refers to molecules. Chiral molecules have the property of rotating the plane of polarization of plane-polarized monochromatic light that is passed through it, this phenomenon is called optical activity; that is why two mirror images of a chiral molecule are called enantiomers or optical isomers. Pairs of enantiomers are often designated as *right-handed* (*R* or *D*) and *left-handed* (*L*). The feature that is most often the cause of chirality in molecules is the presence of an asymmetric carbon atom.¹

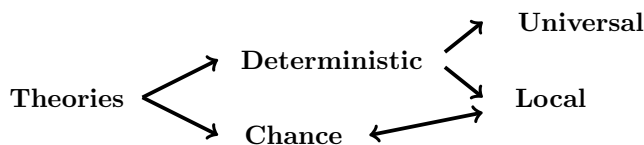
Synthesis of chiral molecules from achiral substrates (in the absence of directing chiral agents) results in the formation of equal amounts of both enantiomers, this is the definition of a racemic mixture, or racemate. By marked contrast, molecules that form the building blocks of life such as amino acids and sugars only exist as a single enantiomer in nature. Indeed, in the artificial synthesis experiments carried out by Miller, for example, only racemates of amino acids were formed [1].

¹An asymmetric carbon atom is a carbon atom that is attached to four different types of atoms or four different groups of atoms.

Homochirality describes a property of some molecules made up of chiral units. A molecule is said to be homochiral if all the molecule's constituent units are of the same chiral form. In biology, homochirality is a common property of amino acids and sugars. It is an empirical fact that there is an absolute chiral imbalance (or mirror symmetry breaking) in all known biological systems, where processes crucial for life such as replication, imply chiral supramolecular structures, sharing the same chiral sign (homochirality). These chiral structures are proteins, composed of amino acids almost exclusively found as the left-handed enantiomers (L); and DNA, and RNA polymers and sugars, with chiral building blocks composed by right-handed (R) monosaccharides.

This raises the fundamental question how these molecules were directed to this asymmetry in the evolution of biological systems. Over the past few decades, experimental observations have provided evidence leading to a variety of theories concerning the origins of biomolecular homochirality [2]. In the absence of a prevalent hypothesis, there is a need to classify the theories proposed so far: if there is a relationship between cause and effect, that is a specific chiral field or influence causing the mirror symmetry breaking, the theory is classified as *deterministic*; otherwise it is classified as a theory based on *chance* (in the sense of randomness) mechanisms.²

Deterministic theories can be further divided into two subgroups: if the initial chiral influence took place in a specific space or time location (averaging zero over large enough areas of observation or periods of time), the theory is classified as *local* deterministic; if the chiral influence is permanent at the time the chiral selection occurred, then it is classified as *universal* deterministic. The classification groups for local deterministic theories and



Scheme 1.1: Classification of the various theories of the origins of biomolecular homochirality.

theories based on chance mechanisms can overlap. Even if an external chiral influence produced the initial chiral imbalance in a deterministic way, the outcome sign could be random since the external chiral influence has its enantiomeric counterpart elsewhere, this is represented in Scheme 1.1.

Deterministic theories evolve through three sequential stages. First, a chiral field or influence creates an initial enantiomeric imbalance in the originally achiral primeval matter. This is the mirror-symmetry breaking step. Once any kind of imbalance has been created, several different mechanisms can operate in further chiral enrichment and perhaps ultimately towards homochirality. This second process is the amplification step. Finally, once a significant enantiomeric enrichment has been produced in a system, the transference of chirality through the entire system is customary. This last step is known as the chiral transmission step.

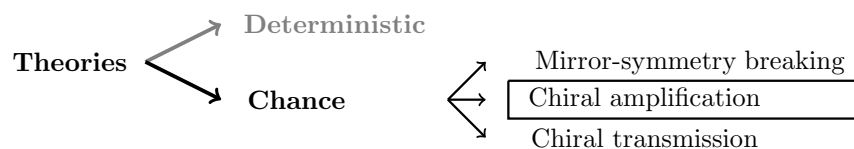
²Another classification for the different theories of the origin of biological homochirality could be made depending on whether life emerged before the enantiodiscrimination step (biotic theories) or afterwards (abiotic theories). Biotic theories claim the coexistence of the two enantiomeric kinds of life (each one having the opposite chiral sign to the other) followed by the extinction of one of them, suggesting that remains of the extinct chirality sign should be found. Since this is not the case, nowadays biotic theories are no longer supported.

In deterministic theories, the enantiomeric imbalance is created due to an external chiral field or influence, and the ultimate sign imprinted in biomolecules will be due to it. So the most important process will be that in which any efficient method can create an initial enantiomeric excess. Thus, for determinist theories, the key process is the mirror-symmetry breaking step.

The study of deterministic theories is beyond the scope of this thesis, but given their prolific nature, it is worth mentioning some examples of mechanisms that can yield an enantiomeric imbalance. Some examples of *local* deterministic theories are: Circular Polarized Light (CPL) on earth as well as in outer space, β -Radiolysis or the magnetochiral effect [3]. The most accepted *universal* deterministic theory is the electroweak interaction.

1.2 Chance Theories

Chance theories are counterintuitive to most chemists, mainly due to the irreproducibility of the experiments that could validate them. The scenarios of Spontaneous Mirror Symmetry Breaking (SMSB) can be included in this category and are the topic of main interest discussed in the following chapters of the present thesis. Chance and deterministic



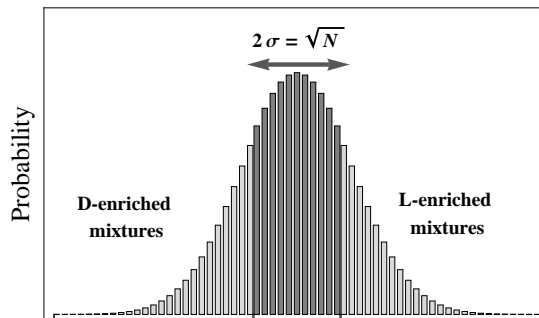
Scheme 1.2: Sequential steps through which chance theories evolve.

theories of the origins of biomolecular homochirality evolve through the same sequential steps or stages: mirror-symmetry breaking followed by chiral amplification and a possible chiral transmission; but the importance of each step is different for each group. In chance theories, the initial enantiomeric excess is inexorable to the system and it is due to stochastic fluctuations within the ideal racemic system. It is based on the assumption that a perfect racemic mixture is chemically impossible to achieve: mirror asymmetric states are in practice unavoidable on purely statistical ground alone [4] (as is further explained in the following sections). In this scenario, there is a need to amplify the initial stochastic chiral imbalance through any efficient mechanism of amplification, so the key process in this case is the amplification step (Scheme 1.2).

Amplification of tiny stochastic imbalances

As already pointed out, the mirror symmetry breaking step is not the key process in theories based on chance mechanisms since there is an imbalance already due to stochastic fluctuations over the ideal racemic system. Consider the racemic state as a macroscopic property described by a binomial distribution, see Fig. 1.1; the experiment of tossing a coin, where the two possible outcomes are the two enantiomers is a good analogy. The discrete probability distribution $P_p(n, N)$ of obtaining n successes out of N Bernoulli trials, where the result of each Bernoulli trial occurs with probability p and the opposite occurs with probability $q = (1 - p)$ is given by Eq.(1.1). As in the experiment of tossing a coin, in this case, we assume both events (L or D) to be equiprobable, $p = q = 1/2$.

Figure 1.1: The statistical distribution of the composition of racemic mixtures can be derived from the binomial distribution, which is characterized by its mean value, $\mu = N/2$ and the standard deviation, $\sigma = \sqrt{N}/2$. About 68% of racemic mixtures (are in dark gray) are within one standard deviation from the mean value, while for the remaining 32% (in lighter gray) the deviation from $N/2$ is greater.



The *mean value* of this distribution is $\mu = Np = N/2$ and the standard deviation is $\sigma = \sqrt{Np(1-p)} = \sqrt{N}/2$, then:

$$P_p(n, N) = \binom{N}{n} p^n (1-p)^{N-n}. \quad (1.1)$$

The racemic state is that in which there are equal amounts of L and D enantiomers of a chiral molecule, in this case, $N/2$ (the case in which N is odd is disregarded for simplicity). The discrete probability distribution $P(N/2, N)$ of having exactly $N/2$ molecules of one chirality and $N/2$ of the other, is given by:

$$P_{1/2}(N/2, N) = \binom{N}{N/2} \left(\frac{1}{2}\right)^{N/2} \left(\frac{1}{2}\right)^{N/2} \approx \sqrt{\frac{2}{\pi N}}, \quad (1.2)$$

using the Stirling's approximation: $N! \approx \sqrt{2\pi N} \left(\frac{N}{e}\right)^N$.

Thus the probability of having exactly the same amount of both enantiomers is inversely proportional to the total number of molecules N . For one mol of a racemic compound, $N = N_A \approx 6.022 \cdot 10^{23}$ molecules, this probability becomes $P_{1/2}(N_A/2, N_A) \approx 10^{-12}$. The most likely mode or configuration is that lying on top of the bell, in this particular case, this mode is the racemic state $(N_A)_L + (N_A)_D$, but the probability to find such a state is so small that we can consider it negligible.

For N large, the binomial distribution approaches a normal (Gaussian) distribution. Following the properties of the normal distribution, about 68% of the values are within one standard deviation from the mean, this is within the interval $\mu \pm \sigma$. Using the values for this particular case, the state of the system will lie within the interval $N_A/2 \pm \sqrt{N_A}/2 = 3.011 \cdot 10^{23} \pm 3.9 \cdot 10^{11}$, about 95% of the values are within two standard deviations from the mean $N_A/2 \pm 2\sqrt{N_A}/2 = 3.011 \cdot 10^{23} \pm 7.8 \cdot 10^{11}$, and about 99.7% lie within three standard deviations from the mean $N_A/2 \pm 3\sqrt{N_A}/2 = 3.011 \cdot 10^{23} \pm 1.7 \cdot 10^{12}$. These results show the great magnitude of the stochastic dispersion for just one mol in the racemic state, leading us to affirm that absolute asymmetric synthesis, even in the absence of chiral physical forces, is in practice unavoidable on statistical grounds alone [4].

That random fluctuations, combined with suitable amplification mechanisms, might ultimately account for biomolecular homochirality in Nature, was clearly reconized more than a century ago by Pearson [5] and it was Mills who first reported a quantitative treatment of the statistical fluctuations that might be responsible for the origin of biomolecular homochirality [6]. The expected number of molecules deviating from the mean value in a sample with N total number of molecules directly depends on N and is given by $\sigma = \sqrt{N}/2$.

It is convenient to define a magnitude describing the degree of chiral imbalance in a given sample, this magnitude is called *enantiomeric excess* and depends on the concentration of both enantiomers as follow:

$$ee = \frac{[L] - [D]}{[L] + [D]} \quad (-1 \leq ee \leq 1). \quad (1.3)$$

Also it is customary to define the same quantity as a percentage:

$$ee(\%) = \frac{[L] - [D]}{[L] + [D]} \cdot 100\%. \quad (1.4)$$

Then, the expected stochastic enantiomeric excess in a sample due to fluctuations (see Fig. 1.1) is given by the product of two standard deviations (where $\sigma = \sqrt{N}/2$) and the probability (0.68%) for the racemic mixture to lie in this interval:

$$ee_{st} = \frac{0.68}{\sqrt{N}}. \quad (1.5)$$

The magnitude of the statistical fluctuation ee_{st} , and so the number of enantiomeric molecules in excess, $ee_{st} \cdot N$, depend on the size of the sample N . Let's consider a feasible scenario to better understand the magnitude of these fluctuations: If we start with a millimolar sample (i.e., customary laboratory amount) of 10^{20} molecules, there is an even chance of obtaining a product containing an excess of $6.7 \cdot 10^9$ molecules, a huge increase of $ee_{st} \cdot N$ that is matched by an equally pronounced decrease of ee_{st} to $6.7 \cdot 10^{-11}$. Even the most powerful measuring device available today is incapable of detecting such a minuscule enantiomeric excess.

Regarding chemical evolution, this obviates the need to invoke chiral physical fields and lends further support to the conviction that homochirality is a "stereochemical imperative" of molecular evolution [7]. Then, the evolution of biological homochirality could be theoretically explained by a model in which a tiny imbalance of one enantiomer was amplified, and the most likely path for this amplification step is by autocatalytic reactions. An autocatalytic chemical reaction is that in which the reaction product is itself a reactive, in other words, a chemical reaction is autocatalytic if the reaction product is itself the catalyst of the reaction. The possible amplification mechanisms that could have lead to the emergence of homochirality are the central point of study of the present thesis.

1.3 Aim and outline of the thesis

The primary research goals of the present work are directed towards understanding the basis of molecular evolution through its origins in prebiotic chemistry, so in this frame work, most of the effort is dedicated to the theoretical and analytical study of the origin of biological homochirality and spontaneous mirror symmetry breaking (SMSB) in chemical systems.

In this context, in Chapter 2, we introduce and study the most basic known autocatalytic system leading to chiral amplification- the Frank model [8]- and its ability to amplify the initial small statistical deviations from the idealized racemic composition. Depending on the conditions, this amplification can be just temporary, and it can be interpreted as a chiral excursion in a dynamic phase space. These chiral excursions can be studied through

a combination of phase space analysis, stability analysis and numerical simulations in order to determine how they depend on whether the system is open, semi-open or closed.

In Chapter 3, the emergence of homochirality in enantioselective autocatalysis for compounds unable to transform according to the Frank-like reaction network is discussed with respect to the controversial limited enantioselectivity (LES) model composed of coupled enantioselective and non-enantioselective autocatalyses, which cannot lead to SMSB in closed systems.

The relationship of the polymerization process with the emergence of chirality is the central point of study in Chapter 4 since biological homochirality of living systems involves large macromolecules [9], thus the ability to amplify (and transfer through the entire system) the small initial enantiomeric excesses due to statistical fluctuations is studied for two different kinetic models of chiral polymerization and copolymerization in systems closed to matter and energy flow, and the results from fitting the copolymerization model to the experimental data [10] on lattice-controlled chiral amplification of oligopeptides are shown.

Finally, in Chapter 5, both a chemical equilibrium model of template-controlled copolymerization and a probabilistic approach are presented for describing the outcome of the experimental induced desymmetrization scenarios recently proposed by Lahav and co-workers [11]; these chiral amplification mechanisms proceed through racemic β -sheet controlled polymerization operative in both surface crystallites as well as in solution. In contrast to the preceding chapters, in this case, the symmetry breaking arises from combinatorics, not from spontaneous (bifurcation) phenomena. These stochastic/statistical/combinatorial effects are not due to the inherent tiny chiral fluctuations present in all real chemical systems but are due rather to the random occlusion of host and guest amino acids by the chiral sites of the template: the mechanisms proposed here work even for ideally racemic mixtures.

Bibliography

- [1] Miller, S. M. *Science* **117**, 528 (1953).
- [2] Guijarro, A. and Yus, M. *The Origin of Chirality in the Molecules of Life* (RSC Publishing, Cambridge, 2009), 1st ed.
- [3] Barron, L. D. *J. Am. Chem. Soc.* **108**, 5539 (1986).
Barron, L. *Molecular Physics* **43**, 1395 (1981).
- [4] Mislow, K. *Collect. of Czech. Chem. C.* **68**, 849 (2003).
- [5] Pearson, K. *Nature* **58**, 495 (1898).
Pearson, K. *Nature* **59**, 30 (1898).
- [6] Mills, W. *J. Soc. Chem. Ind.* **10**, 75 (1932).
- [7] Siegel, J. *Chirality* **10**, 24 (1998).
- [8] Frank, F. *Biochim. Biophys. Acta* **11**, 459 (1953).
- [9] Sandars, P. *Origins Life Evol B* **33**, 575 (2003).
- [10] Zepik, H., Shavit, E., Tang, M., Jensen, T., Kjaer, K., Bolbach, G., Leiserowitz, L., Weissbuch, I. and Lahav, M. *Science* **295**, 1266 (2002).
Weissbuch, I., Zepik, H., Bolbach, G., Shavit, E., Tang, M., Jensen, T., Kjaer, K., Leiserowitz, L. and Lahav, M. *Chem-Eur. J.* **9**, 1782 (2003).
- [11] Nery, J., Bolbach, G., Weissbuch, I. and Lahav, M. *Chem. Eur. J.* **11**, 3039 (2005).

Chapter 2

Frank model

2.1 Introduction

In 1953, Frank described a simple autocatalytic process leading to the amplification of a tiny initial enantiomeric excess which is still widely accepted at the present time: the Frank model. The simple and sufficient model that Frank proposed was based on a chemical substance which is a catalyst for its own production and an anti-catalyst for the production of its optical antimer. Actually, Frank did not propose any set of chemical reactions but of dynamical equations where the concentrations of both enantiomers were denoted as $[n_1]$ and $[n_2]$ respectively, and their rate of increase to be given by the following pair of equations (where time dependence is omitted in the notation for simplicity):

$$\frac{d[n_1]}{dt} = (k_1 - k_2[n_2])[n_1], \quad \frac{d[n_2]}{dt} = (k_1 - k_2[n_1])[n_2], \quad (2.1)$$

where k_1 and k_2 are positive. The interaction between n_1 and n_2 expressed by the second term on the right side of Eq.(2.1) may be interpreted as a lethal interaction (or as a tendency to diminish the reproduction rate of the other).

The analytical solutions for Eq.(2.1) are found to be $[n_1]/[n_2] = [n_1]_0/[n_2]_0 e^{k_2([n_1]_0 - [n_2]_0)(e^{k_1 t} - 1)}$, this ratio increases at a more than exponential rate if $[n_1]_0 - [n_2]_0$ is positive (and viceversa). Every starting conditions different to $[n_1]_0 = [n_2]_0$ lead to one of the asymptotes $[n_1]_0 = 0$ or $[n_2]_0 = 0$. Thus the equality of $[n_1]_0$ and $[n_2]_0$ and so of $[n_1]$ and $[n_2]$ represents a condition of unstable equilibrium, this result depending on the presence of the term representing mutual antagonism.

In order to study the sensitivity of the model described in Eq.(2.1) to the miniscule initial enantiomeric excess, a very dilute initial concentration of a scalemic (non racemic) mixture is used in the following simulations: the initial concentrations of $[n_1]_0 = 10^{-6} + 10^{-15}M$ and $[n_2]_0 = 10^{-6}M$ yielding an initial chiral excess of $ee_0 = 5 \cdot 10^{-10}$. This is actually slightly lower than the excess corresponding to the initial concentrations ($ee_{st} = 6.2 \cdot 10^{-10}$). Setting $k_1 = 5 \cdot 10^{-5} s^{-1}$ and $k_2 = 5 \cdot 10^{-5} s^{-1} M^{-1}$ and numerically integrating over time we obtain the simulations shown in Fig. (2.1a). As shown in Fig. (2.1a), the production of n_1 and n_2 undergoes a bifurcation leading to the exclusive presence of one enantiomeric species after it. The initial enantiomeric excess $ee_0 = 5 \cdot 10^{-8}\%$ is amplified up to its maximum value $ee = 100\%$. If we set up equations representing also unspecific antagonism then:

$$\frac{d[n_1]}{dt} = (k_1 - k_2([n_1] + [n_2]))[n_1], \quad \frac{d[n_2]}{dt} = (k_1 - k_2([n_1] + [n_2]))[n_2], \quad (2.2)$$

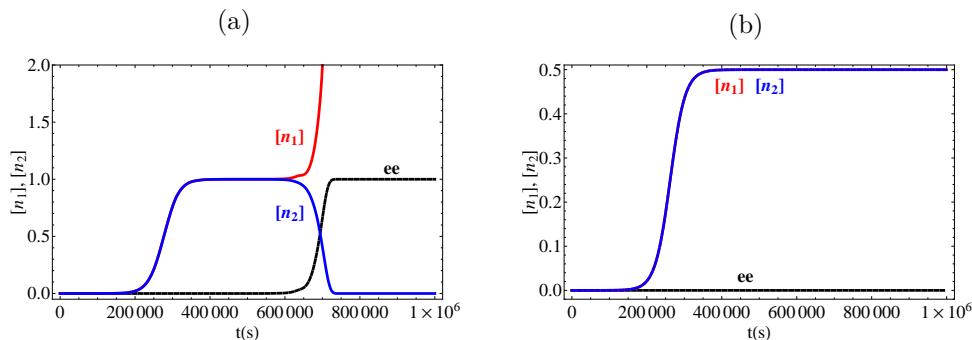
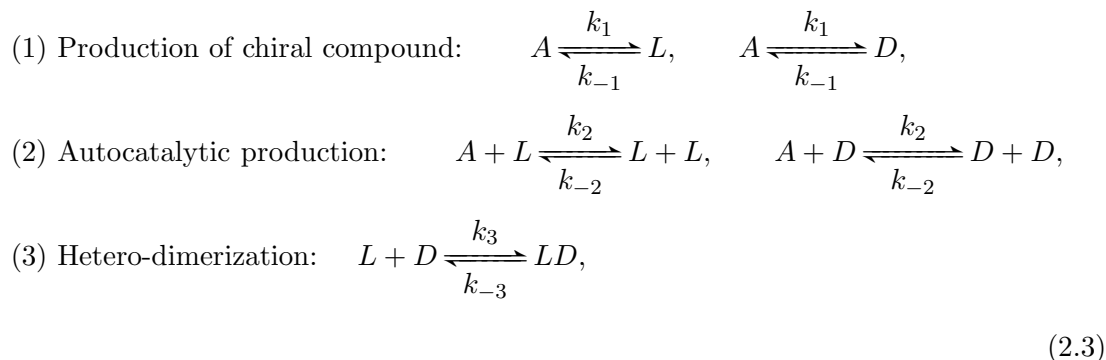


Figure 2.1: Numerical simulation of Frank's model described by (a) Eq.(2.1), for which the initial stochastic enantiomeric excess is amplified up to its maximum value of 100% and (b) Eq.(2.2), for which the initial stochastic enantiomeric excess is not amplified but preserved. Initial conditions for both cases: $[n_1]_0 = 10^{-6} + 10^{-15}M$ and $[n_2]_0 = 10^{-6}M$, and reaction rates: $k_1 = 5 \cdot 10^{-5}s^{-1}, k_2 = 5 \cdot 10^{-5}s^{-1}M^{-1}$. Initial conditions correspond to an $ee_0 = 5 \cdot 10^{-10}$, slightly lower than the calculated stochastic excess, $ee_{st} = 6.2 \cdot 10^{-10}$

with solutions $[n_1] + [n_2] = \frac{k_1/k_2}{1 \pm e^{-k_1 t}}$. When including unspecific antagonism, any initial disproportion is preserved, but not amplified. The model is no longer able to amplify tiny initial enantiomeric excesses. Setting the same parameters as in the example above, the results of the simulation are those shown in Fig(2.1b).

2.2 Reversible Frank model

A much more realistic scenario based on Frank's model was proposed in 1983 by Kondupudi and Nelson [1]. In this model, two kind of achiral precursors, A and B participate in the production of the enantiomeric products L and D in two different ways: direct production and autocatalytic production (we abandon the notation n_1 and n_2 to use a more customary one: L and D), and all the steps are assumed to be reversible. The reaction scheme studied in the following sections consists of a straight non-catalyzed reaction (1), an enantioselective autocatalysis (2) and an heterodimerization step (3). Note that, for the sake of simplicity, A is the only prechiral starting product in this scheme, and L and D are the two enantiomers of the chiral product:



where LD is an achiral product that is usually removed from the system at a rate γ . The elimination of the heterodimer can actually be neglected as long as the mutual inhibition remains irreversible [2]. It is usually to find this model being called "Frank's model" in

the literature, since it is the most widely used adaptation of the original model nowadays. Several versions of the original model have been proposed during the last decades: the two achiral precursors A and B in reactions (1) and (2) [1] can actually be replaced by a single achiral compound, named A [3] [4]; the direct production described in the reaction (1) can be totally neglected compared with the autocatalytic reaction (2) [5–7]; and the mutual inhibition in reaction (3) and the removal of the achiral heterodimer from the system can be modeled by a direct continuous elimination of both chiral compounds L and D from the system [7].

When reversible steps in all the reactions are allowed (and it is assumed to be a closed system) it is capable of [8] (i) amplification of the initially tiny statistical enantiomeric excesses from $ee \sim 10^{-8}\%$ to practically 100%, leading to (ii) long duration chiral *excursions* or chiral pulses away from the racemic state at nearly 100% ee , followed by, (iii) the final approach to the stable racemic state for which $ee = 0$, i.e., mirror symmetry is recovered permanently. To understand this temporary asymmetric amplification is important because the racemization time scale can be much longer than that for the complete conversion of the achiral substrate into enantiomers. In order to better understand these chiral excursions, the Frank model has been studied by combining the information provided by phase plane portraits, numerical simulation and linear stability analysis [4].

The equilibrium constants for the direct production and the autocatalytic reactions in Eq.(2.3) are given by:

$$K_1 = \frac{k_1}{k_{-1}} = \frac{[L]}{[A]} = \frac{[D]}{[A]}, \quad K_2 = \frac{k_2}{k_{-2}} = \frac{[L]^2}{[A][L]} = \frac{[D]^2}{[A][D]}, \quad (2.4)$$

so the system is under the following thermodynamic constraint:

$$\frac{k_1}{k_{-1}} = \frac{k_2}{k_{-2}}. \quad (2.5)$$

2.3 Results

Focusing our attention on chiral excursions, we make a careful distinction between open, semi-open or fully closed systems. These system constraints are crucial for determining both the intermediate and the asymptotic final states of the chemical system. Each case has been studied using information provided by phase plane portraits and linear stability analysis.

For obtaining the fixed points and stability of the system, it is convenient to redefine the differential equations in the variables sums and differences of dimensionless concentrations: $\chi = [\tilde{L}] + [\tilde{D}]$ and $y = [\tilde{L}] - [\tilde{D}]$. The stationary solutions of the kinetic equations for $\frac{d\chi}{d\tau}$ and $\frac{dy}{d\tau}$ (τ is a dimensionless time parameter) correspond to the steady states of the system. Generally, there is an (always unstable) unphysical state called U , a racemic state R and two chiral states Q_{\pm} . The fixed points correspond to the *final* asymptotic state (as $\tau \rightarrow \infty$) solutions of the kinetic model. Substituting $\chi = \chi^* + \delta\chi(t)$, $y = y^* + \delta y(t)$, etc., into the set of differential equations for χ and y , where χ^* , y^* , ..., denotes a fixed point solution, we obtain differential equations for the arbitrary perturbations $\delta\chi(t)$, $\delta y(t)$, ..., about the fixed point. Then, the Jacobian matrix governs the time dependence of these perturbations to first order $O(\delta)$ in the fluctuations. In order to study the stability of the

possible homogeneous solutions (generally R and Q_{\pm}), we calculate the eigenvalues of the Jacobian matrix evaluated at each of the fixed points solutions.

The constraint imposed by Eq.2.5 must be taken into account at all times in order to avoid erroneous and artifactual simulations displaying mirror symmetry breaking. In this aspect, the main point is the issue of kinetic versus thermodynamic control and the reversibility of the chemical reaction steps, or processes, as crucial elements to take into account to avoid misinterpretations.

The key variables throughout for the study of the phase plane portraits are the dimensionless chiral polarization, η , and the total chiral mass of the system, χ :

$$\eta = \frac{[\tilde{L}] - [\tilde{D}]}{[\tilde{L}] + [\tilde{D}]}, \quad \chi = [\tilde{L}] + [\tilde{D}], \quad (2.6)$$

η obeys $-1 \leq \eta \leq 1$, and represents the order parameter for mirror symmetry breaking (see Chapter 1).

In the phase space of the dynamical system defined by the kinetic equations of η and χ there are curves with a special significance. These are the nullclines defined by $\frac{d\chi}{d\tau} = 0$ and $\frac{d\eta}{d\tau} = 0$. The intersections of the curves defined above correspond to the fixed points of the system. The representation of these curves and the study of its morphology help to understand the chiral excursion and predict their appearance.

Fig. 2.2 shows the temporal evolution of the L and D chiral monomer concentrations starting from an extremely dilute total monomer concentration and the very small statistical chiral deviations from the ideal racemic composition for the three different cases.

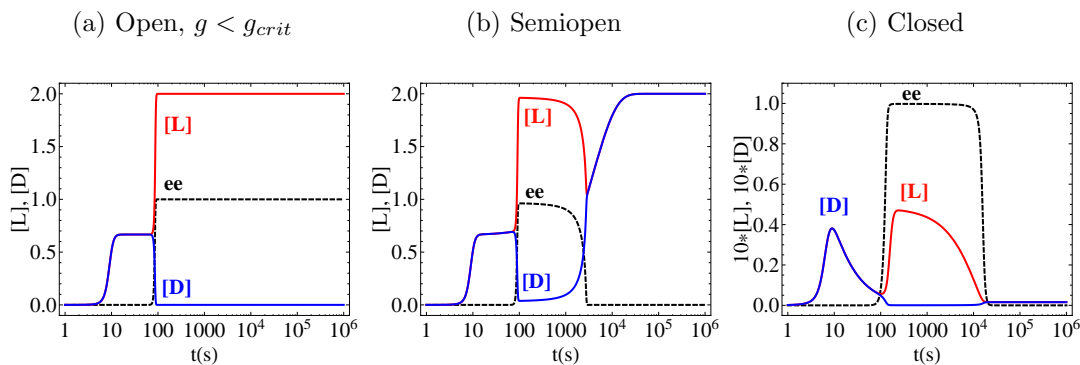


Figure 2.2: Temporal behavior (logarithmic scale) of the individual enantiomer concentrations $[L]$ and $[D]$, in cases of chiral symmetry breaking in the a) open, b) semi-open and c) closed systems (L and D have been amplified by a factor of 10). Initial concentrations: $[L]_0 = (10^{-6} + 10^{-15})M$, $[D]_0 = 10^{-6}M$ ($ee_0 = 5 \cdot 10^{-8}\%$) and $[A] = 1M$. Rate constants: $k_1 = 10^{-4}s^{-1}$, $k_{-1} = 10^{-6}s^{-1}$, $k_2 = 1s^{-1}M^{-1}$ and $k_{-2} = 0.5s^{-1}M^{-1}$. For the case (a) $k_3 = 1s^{-1}M^{-1}$, (b) $k_3 = 1s^{-1}M^{-1}$ and $k_{-3} = 10^{-3}s^{-1}$, and (c) $k_3 = 20s^{-1}M^{-1}$ and $k_{-3} = 10^{-4}s^{-1}$. All the values correspond to the situations in which either the expected final stationary state is the chiral one (open case, $g < g_{crit}$) or a chiral excursion is observed (semi open and closed systems). In figures of simulations, we always display original variables $[L]$, $[D]$, $[LD]$ and $ee = ([L] - [D])/([L] + [D])$, etc. as function of time t . Initial conditions correspond to an $ee_0 = 5 \cdot 10^{-10}$, slightly lower than the intrinsic stochastic excess, $ee_{st} = 6.2 \cdot 10^{-10}$.

Open system

For an open system, the prechiral component A is assumed to be constant and instead of having the heterodimer dissociate into monomers, we set $k_{-3} = 0$ and remove the heterodimer from the system at a fixed rate $\bar{\gamma}$, that is,



The stationary solutions for the open system reveals four steady states: besides an unphysical state, denoted as U , there is a Z_2 pair of chiral solutions Q_{\pm} and a racemic state R . The linear stability analysis determines that the final stable state can be either racemic or chiral, depending on the system parameters' values. The key parameter in this case, is a ratio between reaction rates called g and defined as $g = k_{-2}/k_3$; the chiral state is found to be the stable if $g < g_{crit} \leq 1$, where $g_{crit} = (\sqrt{1+16u} - 1)/8u$ and $u = k_1 k_3 [A] / (k_2 [A] - k_{-1})^2$. Thus, a necessary but not sufficient condition to achieve a final stable chiral state is $k_3 > k_{-2}$. That is, the chiral state is stable if and only if the heterochiral interactions are favored over the homochiral ones. Fig. 2.2a shows the the temporal evolution of $[L]$, $[D]$ and η for the case in which permanent symmetry breaking is expected.

The direct monomer production step (k_1) tends to racemize the system leading to final ee values strictly less than unity: the monomer production step reduces the range of parameters' values for which stable mirror symmetry breaking can occur, and the chiral solutions are no longer 100% chiral.

Are there chiral excursions found in the open system model? A chiral excursion holds when the enantiomeric excess η departs from a small initial value, evolves to some maximum absolute value and then decays to the final value of zero. The nullclines for this case show, depending on the parameters's values, three possible intersections, R and Q_{\pm} for $g < g_{crit}$ (Fig. 2.3a, corresponding to Fig. 2.2a) or a single intersection R for $g > g_{crit}$ (indicating that the racemic state is the only possible solution, Fig. 2.3b); in qualitative agreement with the stability analysis. We set $g > g_{crit}$ to ensure a final racemic state (see

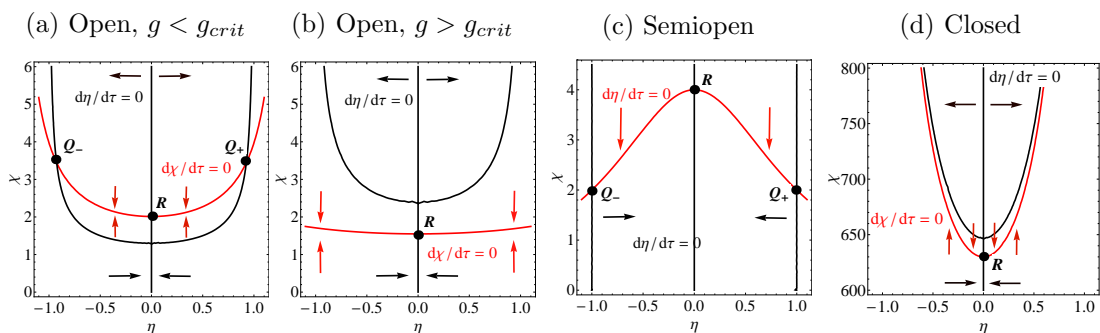


Figure 2.3: Nullclines for the (a),(b) open, (c) semi-open and (d) closed systems (for more information on the parameters' values, see Ref. [4]). The $d\eta/d\tau = 0$ and $d\chi/d\tau = 0$ nullclines are plotted in black and red, respectively. The black (red) arrows indicate the regions of phase space where η (χ) increases or decreases.

Fig. 2.3b) but then we find no numerical evidence for such temporary chiral excursions.

This can be understood qualitatively from inspection of the nullclines: the initial conditions (dilute monomer concentration and statistical chiral fluctuation) corresponds to a initial point located at tiny values of both χ and η . In this situation, the system approaches the single possible intersection of curves, R , and it is impossible for the chiral excess to increase, not even temporarily. On the other hand, if we set the parameters to ensure a final chiral state, $g < g_{crit}$ (see Fig. 2.3a), the system approaches the racemic state R , but before its arrival, it is attracted to one of the other possible two chiral fixed points where it stays forever, provided the system is maintained out of equilibrium. The chiral symmetry is permanently broken, and there is no excursion such as we have defined it.

Semiopen system

For the semi-open case, we do not remove the heterodimer from the system (and allow the back reaction to homodimers), although we still keep $[A]$ constant.

In this case, we find negative concentrations in the chiral states Q_{\pm} , so clearly, for this scheme, the chiral states are also unphysical states. Even if the linear stability analysis implies extremely long algebraic expressions, an inequality analysis over the racemic state shows that two of the three corresponding eigenvalues are negative and the other one is null, concluding that the racemic state R is *marginally* stable (see Fig. 2.2b). The final outcome will always be the racemic state. There is no stable mirror symmetry broken solution when the heterodimer dissociates back into the chiral monomers. Nevertheless, the system can have temporary chiral excursions (see Fig. 2.2b and Fig. 2.3c).

Closed system

There is no flow of material into or out of the system, thus $[A]$ is not constant in this situation.

The dimensionless concentrations are subject to the constraint $[\tilde{A}] = [\tilde{C}] - ([\tilde{L}] + [\tilde{D}]) - 2[\tilde{LD}]$, where $[\tilde{C}] = (k_3/k_1)[C]$ and $[C]$ is the total initial concentration, being constant in time.

As before, we find unphysical chiral solutions Q_{\pm} since they imply negative monomer concentrations ($\chi < 0$). Thus, the only physically acceptable solution is the racemic one R , and this is (at least marginally) stable.

The nullclines for this case, show a single intersection R indicating that the racemic state is the only possible solution, in qualitative agreement with the stability analysis. The final outcome will always be the racemic state. There is no stable mirror symmetry broken solution for a closed system (to matter and energy). Nevertheless, the system can have temporary chiral excursions (see Fig. 2.2c and Fig. 2.3d).

2.4 Conclusions

- Enantioselective autocatalysis [2], when coupled to other reactions implying chiral catalysts, may lead to SMSB. In the Frank model [5], enantioselective autocatalysis amplifies the chiral fluctuations by the coupling with a "mutual inhibition" reaction that determines (by removing the chiral products/catalysts in racemic composition) the increase of the *ee* of the catalyst available for autocatalysis.

- Temporary chiral excursions are observed for closed and semi-open systems and can be explained through phase space analysis, stability analysis and numerical simulations. Such chiral excursions may be experimentally observed and could be mistaken for a transition to a chiral state.

In the closed system (neither inflow nor outflow), total mass is conserved and temporary symmetry breaking can occur but never permanent symmetry breaking: SMSB is a kinetically controlled emergence of chirality, however, the system is resistant to the eventual racemization that occurs much later than the time necessary to exhaust the initial achiral products.

For the semi-open case, LD is not removed and we allow for its dissociation into chiral monomers. There is no mass balance but temporary symmetry breaking can arise.

- A recent kinetic analysis of the Frank model in closed systems applied to the Soai reaction [9] indicates that in actual chemical scenarios, reaction networks that exhibit SMSB are extremely sensitive to chiral inductions due to the presence of inherent tiny initial enantiomeric excesses [8].

Finally, Mauksch and Tsogoeva have also previously indicated that chirality could appear as the result of a temporary asymmetric amplification [10, 11].

Bibliography

- [1] Kondepudi, D. and G.W., N. *Phys. Rev. Lett.* **50**, 1023 (1983).
- [2] Plasson, R., Kondepudi, D. K., Bersini, H., Commeyras, A. and Asakura, K. *Chirality* **19**, 589 (2007).
- [3] Saito, Y. and Hyuga, H. *J. Phys. Soc. JPN* **74**, 1629 (2005).
- [4] **Blanco, Celia**, Stich, M. and Hochberg, D. Temporary mirror symmetry breaking and chiral excursions in open and closed systems. *Chem. Phys. Lett.* **505**, 140–147 (2011).
- [5] Frank, F. *Biochim. Biophys. Acta* **11**, 459 (1953).
- [6] Saito, Y. and Hyuga, H. *J. Phys. Soc. JPN* **73**, 1685 (2004).
- [7] Iwamoto, K. *Phys. Chem. Chem. Phys.* **5**, 3616 (2003).
- [8] Crusats, J., Hochberg, D., Moyano, A. and Ribó, J. M. *ChemPhysChem* **10**, 2123 (2009).
- [9] Soai, K., Shibata, K. T., Morioka, H. and Choji, K. *Nature* **378**, 767 (1995).
- [10] Mauksch, M., Tsogoeva, S. B., Wei, S. and Martynova, I. M. *Chirality* **19**, 816 (2007).
- [11] Mauksch, M. and Tsogoeva, S. B. *Chem. Phys. Phys. Chem.* **9**, 2359 (2008).

Chapter 3

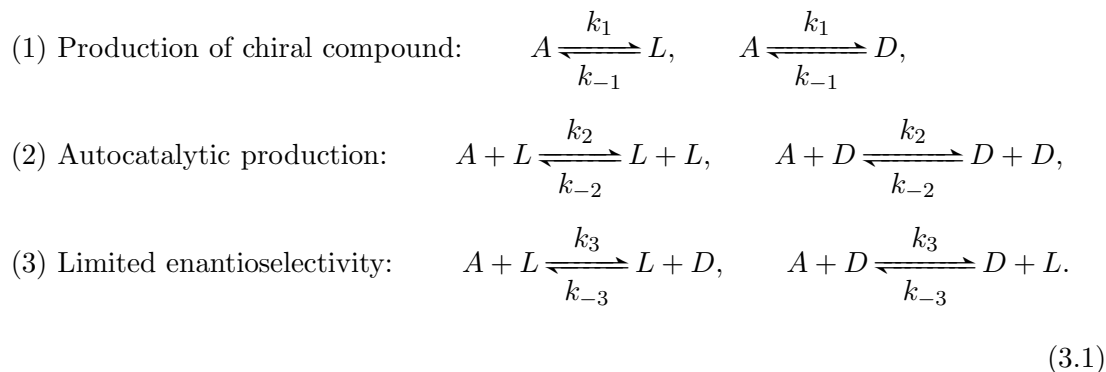
Limited enantioselective (LES) autocatalytic model

3.1 Introduction

Enantioselective autocatalysis is the key reaction that may lead to SMSB on coupling to reactions that amplify the enantiomeric excess (ee) generated by statistical fluctuations about the ideal racemic composition (or induced by small chiral polarizations). Clearly, for a system in conditions that allow a chemical thermodynamic equilibrium to be reached (i.e. a closed system with a uniform distribution of matter, temperature, and energy), the final outcome must be the racemic state. However, in a system with a nonuniform energy distribution (e.g. energy absorption by only some of the species of the system, or open to matter exchange with the surroundings), the final stable stationary state may be chiral, as has been already seen for the Frank model of SMSB in open systems.

In a Frank-like network a necessary condition for SMSB is that the heterochiral interaction between products/catalysts is favored compared to the homochiral interaction [1]. This is the case in the majority of chiral organic compounds, as can be inferred from the number of chiral compounds that crystallize as racemic crystals compared to those yielding a racemic mixture of enantiopure crystals (racemic conglomerates) [2]. However, this last group includes significant compounds in prebiotic chemistry, as for example, several amino acids. The emergence of chirality in enantioselective autocatalysis for compounds which do not follow Frank-like schemes is investigated here for the limited enantioselectivity (LES) model composed of coupled enantioselective and non-enantioselective autocatalyses.

The LES model [3] is composed of an enantioselective autocatalytic reaction coupled to the corresponding nonenantioselective autocatalysis, and was proposed originally as a SMSB alternative to the Frank model [4]. The basic model [3] is defined by the following chemical transformations.



In both models (Frank and LES) the enantioselective autocatalysis in reaction (2) is coupled to a reaction leading to the decrease of the racemic composition of chiral catalysts: in Frank because of the "mutual inhibition" between enantiomers and in LES because of the reverse reaction of the non-enantioselective autocatalysis (3). In contrast to the Frank model, LES is able to account for two important facts: namely, (i) the enantioselectivity of any chiral catalyst is limited because of reaction (3), and (ii) the kinetic link between mirror conjugate processes arises from the reversibility of the catalytic stage [3]. The inverse reaction of the non-enantioselective autocatalysis (reaction (3)) substitutes for the mutual inhibition reaction in the Frank model or formation of the heterodimer ($L + D \rightarrow P$).

The equilibrium constants for the direct production and the autocatalytic reactions in this case are given by:

$$K_1 = \frac{k_1}{k_{-1}} = \frac{[L]}{[A]} = \frac{[D]}{[A]}, K_2 = \frac{k_2}{k_{-2}} = \frac{[L]^2}{[A][L]} = \frac{[D]^2}{[A][D]}, K_3 = \frac{k_3}{k_{-3}} = \frac{[L][D]}{[A][L]} = \frac{[D][L]}{[A][D]}, \quad (3.2)$$

so the system is under the following thermodynamic constraint:

$$\frac{k_1}{k_{-1}} = \frac{k_2}{k_{-2}} = \frac{k_3}{k_{-3}}. \quad (3.3)$$

For a closed system, the dimensionless concentrations are subject to the constraint $[\tilde{A}] = [\tilde{C}] - ([\tilde{L}] + [\tilde{D}])$, where $[\tilde{C}] = (k_3/k_1)[C]$ and $[C]$ is the total initial concentration, being constant in time.

The linear stability study is performed by following the same procedures explained in Chapter 2: to obtain the stationary solutions, we redefine the differential equations in the variables sums and differences of dimensionless chiral concentrations, $\chi = [\tilde{L}] + [\tilde{D}]$ and $y = [\tilde{L}] - [\tilde{D}]$, and calculate the steady states of $\frac{d\chi}{d\tau}$ and $\frac{dy}{d\tau}$. The system has four steady states: besides an unphysical state, denoted as U , there is a Z_2 pair of chiral solutions Q_{\pm} , and a racemic state R . The stability of the asymptotic racemic solution is governed by the algebraic sign of the associated eigenvalues. Performing an inequality analysis over the racemic state it is found that the condition for SMSB in LES may occur when, in addition to $k_2 \gg k_1$ and $k_2 > k_3$, the condition $g < g_{crit}$ where $g = k_{-2}/k_{-3}$ and $g_{crit} \approx (1 - k_3/k_2)/(1 + 3k_3/k_2)$ holds. Obviously this cannot be achieved when the thermodynamic constraint of equation 3.3 is fulfilled (see Ref. [5]).

Previous reports had claimed SMSB in this model; contradictory reports concerning this were consequence [6] of the use of a set of reaction rate constants which do not fulfill the principle of detailed balance.

A new scenario for SMSB in compounds for which the homochiral interactions are more favored than the heterochiral ones, i.e. for reactions that cannot follow a Frank-like model or scheme, is that of the dramatic experimental reports on the deracemization of racemic mixtures of crystals and on the crystallization from boiling solutions [7], [8], [9], [10], [11], [12]. This probably also occurs for other phase transitions, as indicated by a recent example on sublimations [13].

In spite of some controversy about the actual mechanisms acting in this SMSB, the experimental reports all agree that the final chiral state is a stationary state; for instance a mechano-stationary state in the case of wet grinding of racemic crystal mixtures [9],

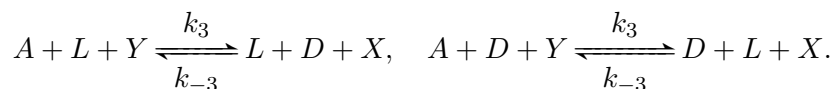
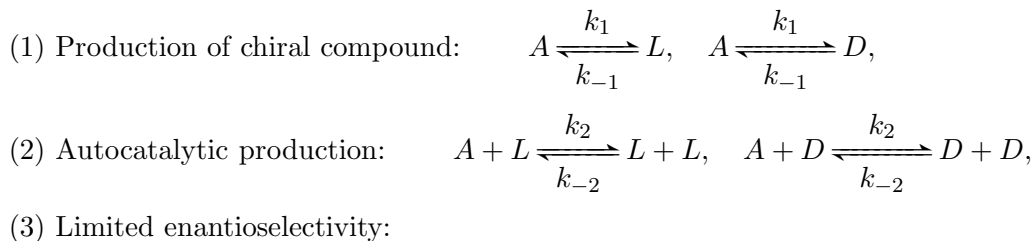
[10], [11] and due to the presence of temperature gradients (system without a uniform temperature distribution) in the case of deracemizations and crystallizations in boiling solutions [7], [8]. Only when the final conditions are those of constant energy input to the system and constant entropy production output to the surroundings, can the system be maintained away from the racemic state in a stationary manner. Note that this raises the question if the LES model could lead to SMSB as a stationary stable state when the system possesses a non-uniform temperature distribution, as is the case of the recent reports of the deracemization of racemic conglomerates. In the following section, we will discuss the possibility and necessary conditions for SMSB in the LES autocatalytic model driven by external reagents and under a temperature gradient.

3.2 Results

LES cycle driven by an external reagent

Following Section 3.1, the LES model in closed systems, and under experimental conditions able to achieve chemical equilibrium, can give rise to neither spontaneous mirror symmetry breaking (SMSB) nor kinetic chiral amplifications. However, it is able to lead to SMSB, as a stationary final state, in thermodynamic scenarios involving a cyclic network if the reverse reaction of the nonenantioselective autocatalysis, which gives limited inhibition on the racemic mixture, is driven by an external reagent, that is, in conditions that keep the system out of chemical equilibrium.

Herein, it is of interest to consider the LES model considering the nonenantioselective autocatalysis driven by a constant concentration of external reactants, X and Y in the reaction (3) as follows:



that is, an open system with X and Y matter exchange with the surroundings.¹

The stability analysis previously carried out for LES in closed systems can be applied here by formally substituting $k_3 \rightarrow [Y]k_3$ and $k_{-3} \rightarrow [X]k_{-3}$ in the relevant expressions for the static solutions and eigenvalues (for a detailed explanation see Ref. [14]). Note that the principle of microreversibility requires:

$$\frac{k_1}{k_{-1}} = \frac{k_2}{k_{-2}} \tag{3.4}$$

And the dimensionless total system concentration is defined by $\tilde{C} = \tilde{A} + \tilde{L} + \tilde{D}$.

¹Notice that, regarding the substrate and the final products, the reaction network is a cyclic one with permanent consumption and production of X and Y , respectively (or of Y and X depending on the reaction flow direction in the cycle).

After obtaining the temporal asymptotic R and Q_{\pm} stationary solutions, it is sufficient to establish the stability criteria of the final state in terms of the racemic solution R alone (algebraic expressions for Q_{\pm} are usually much more complex). The stability of time asymptotic racemic solution is governed by the algebraic signs of the associated pair of eigenvalues: stability in state R requires both of these eigenvalues to be negative, $\lambda_{1,2}(R) < 0$. The first expression $\lambda_1(R) < 0$ is always negative, whereas the second $\lambda_2(R)$ can be positive or negative depending on the values of the parameters. The critical parameter g in this case is defined as $g = \frac{k_{-2}}{[X]k_{-3}}$, and the racemic state is stable ($\lambda_2(R) < 0$) if and only if $g > g_{crit}$, where:

$$g_{crit}(u) = -\frac{1}{8\tilde{C}} \left\{ 4\tilde{C}(h-r) + \tilde{C}^2 h^2 + 2\tilde{C}^2 hr - 3\tilde{C}^2 r^2 + 4u - 2\tilde{C}u(h+r) + u^2 \right. \\ \left. + (\tilde{C}(r-h) + u) \sqrt{\tilde{C}^2(h+3r)^2 + 8\tilde{C}(2+h+3r) - 2\tilde{C}(h+3r)u + (4+u)^2} \right\}. \quad (3.5)$$

The dimensionless parameters appearing on Eq.(3.5) are defined as:

$$u = \frac{k_{-1}}{k_1}, \quad h = \frac{k_2}{[X]k_{-3}}, \quad r = \frac{[Y]k_3}{[X]k_{-3}}. \quad (3.6)$$

Since $g > 0$ always, studying the sign of g_{crit} we can obtain conditions for the final asymptotic states, g_{crit} is found to be real and positive if and only if $h > r$ and

$$C > C_{min} = \frac{u+2ru}{h-r} + 2\sqrt{\frac{u(h-r)+r^2u^2}{(h-r)^2}}. \quad (3.7)$$

Thus, if either of these two conditions is not satisfied (either $h \leq r$ or $C \leq C_{min}$), then $g_{crit} < 0$ and because g is always positive, there can be no values of g satisfying $g < g_{crit}$, and consequently the only stable state will be R , the racemic one. So, we can conclude that the racemic state will be stable under two possible conditions:

$$R \text{ is stable if } \begin{cases} h \leq r & \text{or} & C \leq C_{min} = \frac{u+2ru}{h-r} + 2\sqrt{\frac{u(h-r)+r^2u^2}{(h-r)^2}} \\ h > r & \text{and} & C > C_{min} \quad \text{and} \quad g > g_{crit}(u). \end{cases} \quad (3.8)$$

Obviously, an unstable racemic state implies a final stable chiral outcome Q_{\pm} , so we can write the only condition that must be fulfilled for the chiral states Q_{\pm} to be stable as:

$$Q_{\pm} \text{ is stable if } h > r \quad \text{and} \quad C > C_{min} \quad \text{and} \quad g < g_{crit}(u). \quad (3.9)$$

All simulations performed by numerical integration show, at long reaction times, the achievement of either a racemic or a chiral final state in accord with the predictions summarized in Eq.(3.8) and Eq.(3.9), respectively. The numerical simulations (see below) reveal a clear picture of the chiral final state in complete accord with the theoretical bifurcation scenario depicted by the equations obtained. After the bifurcation point giving the transition from the racemic state to the chiral state, the racemic solution is a metastable state with respect to the more stable degenerate chiral solutions. Furthermore, the calculated χ values for all the racemic and metastable chiral states and of the chiral solutions agree at the level of the numerical error with the values obtained in the numerical simulations.

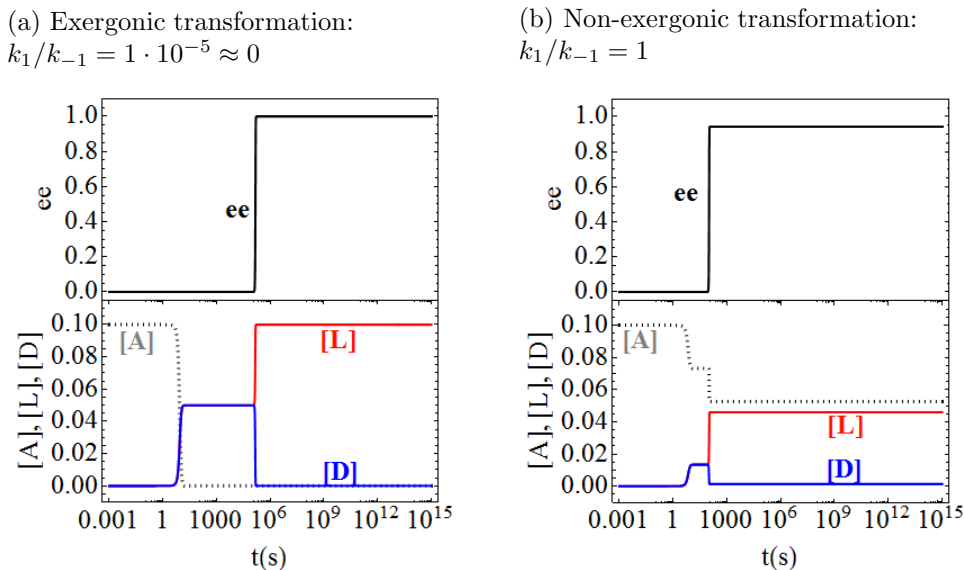


Figure 3.1: Examples of the time evolution of the LES reaction network to achieve the final chiral stable state ($g > g_{crit}$), corresponding to systems showing an enantioselective autocatalysis much faster than the nonautocatalytic reaction ($k_2 \gg k_1$), for (a) an exergonic reaction ($k_{-1}/k_1 \approx 0$) and (b) a nonexergonic reaction ($u = 1$). The final state obeys the criteria of Eq.(3.8) and Eq.(3.9), that is, when $h > r$ or $C > C_{min}$, and the final concentrations agree with those calculated from the asymptotic solutions ($t \rightarrow \infty$). Initial conditions: $[A]_0 = 0.1M$, $[L]_0 = 10^{-10} + 10^{-20}M$, $[D]_0 = 10^{-10}M$ and $[Y]_0 = [X]_0 = 0.01M$. Reaction rates: (a) $k_1 = 10^{-4}s^{-1}$, $k_{-1} = 10^{-9}s^{-1}$, $k_2 = 10s^{-1}M^{-1}$, $k_{-2} = 10^{-4}s^{-1}M^{-1}$, $k_3 = 10s^{-1}M^{-2}$, $k_{-3} = 0.5s^{-1}M^{-1}$; (b) $k_1 = 10^{-4}s^{-1}$, $k_{-1} = 10^{-4}s^{-1}$, $k_2 = 10s^{-1}M^{-1}$, $k_{-2} = 10s^{-1}M^{-1}$, $k_3 = 100s^{-1}M^{-2}$ and $k_{-3} = 5000s^{-1}M^{-2}$. Initial conditions correspond to an $ee_0 = 5 \cdot 10^{-11}$, considerably lower than the calculated stochastic excess, $ee_{st} = 6.2 \cdot 10^{-8}$.

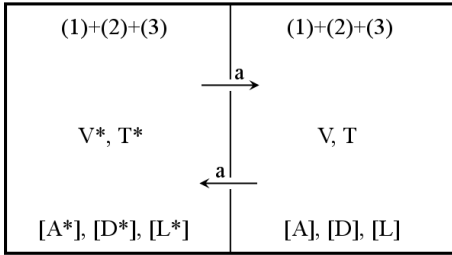
Fig.3.1 shows two examples of the time evolution for the LES reaction network under conditions leading to a final chiral stable state. For the same reaction parameters, the change of k_{-3} (or X) may lead to chiral or racemic final states. This simple change of one of the ten reaction parameters leads to the change of three (h , r , and C) of the five dimensionless parameters. The first example (3.1a) corresponds to an exergonic reaction ($k_1/k_{-1} = k_2/k_{-2} = 1 \cdot 10^{-5}$) and the second example (3.1b) illustrates a nonexergonic transformation ($k_1/k_{-1} = k_2/k_{-2} = 1$). Both examples show how the deracemization occurs later than that of the transformation of A to D and L .

Conversion of the cyclic LES model into a Frank model. If reaction (3) is replaced by the reaction $2A + Y = D + L + X$, the cyclic LES system transforms into a cyclic Frank-like reaction network, which can also be solved exactly for the asymptotic analytic solutions. For all simulated examples [14] for the same reaction parameters but using the above mutual inhibition reaction instead of reaction (3), no significant differences were observed between both models. All cases exhibit the transition between the chiral and racemic states for the same reaction parameters and lead to similar final concentrations. This suggests that the importance of the inverse reaction in (3) and of the (forward) mutual inhibition in the Frank model is due to the decrease of racemic composition regardless of whether or not it leads to the partial or total destruction of a racemic mixture. To bring these results closer to applied chemistry topics and to the emergence of chirality

in prebiotic chemistry, a speculative but plausible scenario is discussed in Ref. [14]. We assume as possible a cyclic reaction network composed of the Strecker amino acid synthesis and the Strecker amino acid degradation and propose stereoselectivity requirements for a reasonable applied chemistry scenario.

LES in a temperature gradient

Consider now the LES scheme described by Eq. 3.1 in a temperature gradient: this permits the reaction rates to vary spatially in the system, from one place to another, and might provide a way to achieve mirror symmetry breaking. The reverse reaction of (3) in one region (compartment) could be faster than the reverse of (2) in the other.



Scheme 3.1: LES in two compartments with volumes V and V^* each held at different temperatures $T^* > T$ and interconnected by an internal flow a of material.

$V([A] + \chi) + V^*([A^*] + \chi^*) = C$, where C is the total conserved mass in the complete two-compartment system. In Ref. [16] it is demonstrated that a temperature gradient and internal flow are by themselves, not enough to produce a bifurcation.

Certain thermodynamic relationships hold among the reaction rates in both compartments. These will be used to prove that SMSB is also impossible for the scheme presented in Section 3.1 in a background temperature gradient. Following this demonstration, we then introduce the variant of LES that can and does lead to SMSB.

From Arrhenius-Eyring, the forward (and reverse) reaction rates for reaction i at temperatures T^* (idem for T , $T < T^*$) are [17]

$$k_i^* = \left(\frac{k_B T^*}{h} \right) e^{-\frac{\Delta G_i(T^*)}{RT^*}}, \quad k_{-i}^* = \left(\frac{k_B T^*}{h} \right) e^{-\frac{\Delta G_{-i}(T^*)}{RT^*}}, \quad (3.10)$$

where $\Delta G_i(T) = \Delta H_i - T\Delta S_i$ denotes the difference in free energy between the activated state (transition state) and the reactants, while $\Delta G_{-i}(T) = \Delta H_{-i} - T\Delta S_{-i}$ is the free energy difference between the activated state (transition state) and the products; H and S denote the enthalpy and entropy, respectively. Clearly, once the values of the k_i are chosen for the reference compartment at T , we are not free to *independently* choose the reaction rates k_i^* at the higher temperature T^* .

The fundamental microreversibility condition in Eq.(3.3) ($\frac{k_i}{k_{-i}} = K(T)$), together with Arrhenius-Eyring implies $\Delta G_{-i} - \Delta G_i \equiv \Delta\Delta G$, that is, the difference of the free energy

And then mixing could bring the hot and cold material into contact. We model this by extending the basic LES scheme in 3.1 to a closed two-compartment system with volumes V, V^* ; each compartment held at a uniform temperature, $T^* > T$, and internally coupled by a constant internal flow or recycling a . The concentrations, and reaction rates for the second compartment are labeled by an asterisk, see Fig. 3.1.

The corresponding rate equations for the two-compartment system have been derived in [15]. As before, for setting up a stability analysis it is convenient to employ the sums and differences of the concentrations $\chi = [L] + [D]$, $y = [L] - [D]$, $\chi^* = [L^*] + [D^*]$ and $y^* = [L^*] - [D^*]$. The rate equations satisfy the constant mass constraint:

differences $\Delta\Delta G$ must be *independent* of i , that is, independent of the specific i th reaction. This implies that the individual double differences in enthalpy and in entropy must also be independent of reaction i , so we must also have $\Delta\Delta H = (\Delta H_{-i} - \Delta H_i)$ and $\Delta\Delta S = (\Delta S_{-i} - \Delta S_i)$.

This also gives us an expression for calculating K , as $K(T) = e^{\frac{\Delta\Delta H}{RT}} e^{-\frac{\Delta\Delta S}{R}}$. The inverse reaction rates are obtained through the constraint Eq.(3.3) as $k_{-i} = \frac{k_i}{K(T)}$.

We also note that if the constraint Eq.(3.3) is satisfied at one specific temperature T , then it will automatically hold at all others, that is

$$\frac{k_i^*}{k_{-i}^*} = K(T^*) = e^{\frac{\Delta\Delta H}{RT^*}} e^{-\frac{\Delta\Delta S}{R}}, \quad (1 \leq i \leq 3). \quad (3.11)$$

The ratio of the equilibrium constants is given by

$$\frac{K(T^*)}{K(T)} = \exp\left(\frac{\Delta\Delta H}{R}\left(\frac{1}{T^*} - \frac{1}{T}\right)\right). \quad (3.12)$$

In view of the gradient $T < T^*$, the putative condition, which could conceivably lead to symmetry breaking in the limit of small values of a ²:

$$k_{-2} < k_{-3}, \quad k_2 > k_3 \quad \text{at } T \quad \text{and} \quad k_{-2}^* < k_{-3}^*, \quad k_2^* > k_3^* \quad \text{at } T^*, \quad (3.13)$$

is incompatible with the constraints in Eqs. (3.3,3.11). This condition (3.13) is inspired by the observation that for $a \rightarrow 0$, the two compartments are practically isolated from each other and can be treated as approximately independent. These could be then the necessary conditions for obtaining an unstable racemic solution in each compartment. But they are incompatible with microreversibility.

Secondly, the analysis in Section 3.1 suggests that SMSB might occur when the inverse reaction of (3) in one region is *faster* than the inverse reaction of (2) in the other region. Taking microreversibility into account, the only way this might be achieved is, for example, by arranging for

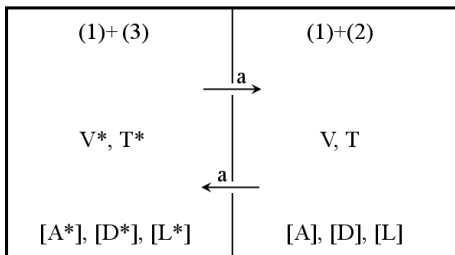
$$k_{-2} > k_{-3}, \quad k_2 > k_3 \quad \text{at } T \quad \text{and} \quad k_{-2}^* < k_{-3}^*, \quad k_2^* < k_3^* \quad \text{at } T^*, \quad (3.14)$$

but this is forbidden by virtue of Eq.(3.12), which is satisfied by the ratio of the equilibrium constants. So a temperature gradient and internal flow are by themselves, not enough to produce a bifurcation. Actually, no spatially varying temperature profile is sufficient, as can be seen by partitioning the closed system into a number of sufficiently small regions within which the local temperature is approximately uniform.

LES in a temperature gradient and compartmentalized autocatalysis (2) and (3)

The condition $k_{-3}^* > k_{-2}$ and $k_3^* < k_2$ is possible thermodynamically in a thought experiment that assumes, in addition to the presence of a temperature gradient, the separation of the autocatalysis reactions in the different temperature regions of the system; the nonenantioselective autocatalysis is confined to the compartment at higher temperature (T^*), whereas the enantioselective autocatalysis is localized within the compartment at the lower temperature (T), see Scheme 3.2.

²The algebraic intricacy of the model is already such that we are unable to obtain *useful* and manageable analytic closed form expressions for the conditions leading to the instability of the racemic solution, and the situation is even worse for obtaining analytic information regarding the possible stationary chiral solutions. We appeal instead to chemically inspired conjectures that can be tested numerically for coherence and compatibility with microreversibility.



Scheme 3.2: Limited enantioselectivity (LES) in two compartments with volumes V and V^* each held at different temperatures $T^* > T$ and interconnected by an internal flow a of material and enantioselective autocatalysis separated from non-enantioselective autocatalysis.

evolution of the species concentrations can be qualitatively described to be composed by three stages, as follows: (i) evolution of the concentrations of A , D , and L to equalize their concentrations in both compartments (can not be seen in the examples of Fig. 3.2 because of the equal initial concentrations in both compartments), followed by (ii) conversion of the species toward the racemic steady state and finally, (i) amplification of this persistently small ee to a high final ee in the case of SMSB (bifurcation), or else racemization in the case of a racemic final state. When the internal flow value a is very small, then the concentration equalization between both compartments occurs later than the transformation of the species, and then racemization dominates over the enantioselective amplification for any initial ee . The LES system of Fig. 3.2 can lead to SMSB at a very low total concentration of the reactants ($[A]_0 + [D]_0 + [L]_0$). Furthermore, the achievement of the bifurcation leading to SMSB, when the condition $k_{-3}^* > k_{-2}$ and $k_3^* < k_2$ is fulfilled, occurs below a critical value of the total concentration; above this critical concentration, the system evolves to the racemic state, and below it to the chiral state. Fig. 3.2 shows an example of how the increase of concentration values -from micromolar to millimolar order- leads from SMSB to a racemic final stationary state (compare middle column and right columns of Fig. 3.2). This critical concentration value depends also on the value of the exchanging flow a with respect to the compartment volumes V and V^* .

A phase and stability analysis of this system would in principle establish the mathematical dependence of all the system variables in an expression for its critical parameter; even for this case, obtaining exact expressions for the conditions that must be satisfied to obtain SMSB is not possible, it is useful to map out the regions in chemical parameter space in which the racemic state is unstable and bifurcates to chiral solutions (see Ref. [16]).

3.3 Conclusions

- The LES model may lead to SMSB in a cyclic network with uniform temperature distribution if the reverse reaction of the nonenantioselective autocatalysis is driven by an external reagent (i.e. in conditions that keep the system out of chemical equilibrium).

For specific values of the reaction network parameters (concentrations and rate constants at T and T^*) and of the system parameters (V , V^* , and a), the final stable state can be a chiral stationary state of high ee . This occurs when the autocatalysis (2) is more effective than the uncatalyzed plain reaction (1) ($k_2 \gg k_1$ and $k_2^* \gg k_1^*$) and $k_{-3}^* > k_{-2}$ while $k_3^* < k_2$. However, these conditions in the two-compartment system are probably necessary but not sufficient.

Fig. 3.2 displays characteristic examples of such simulations obtained by numerical integration of the corresponding differential equations (see Ref. [16]) modified assuming the presence or absence of the transformations (2) or (3) in each compartment. Reasonable reaction parameters were used assuming a very slow uncatalyzed reaction (1) and good autocatalysis (2) and (3). In the case of SMSB, the

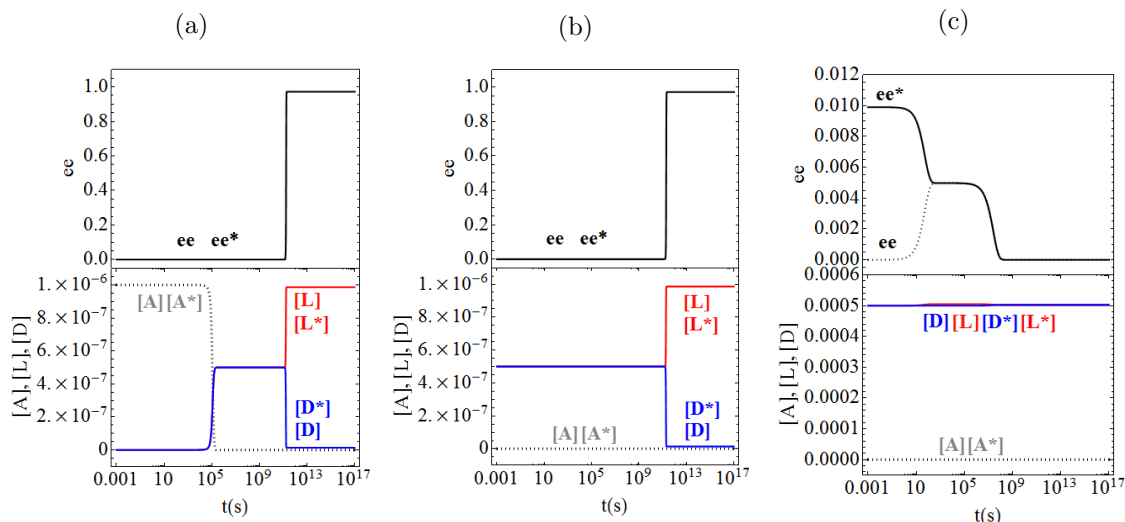


Figure 3.2: Simulations of the LES model in two compartments at different temperature (T and T^*) and enantioselective autocatalysis separated from non-enantioselective autocatalysis under the condition $k_{-3}^* > k_{-2}$ and $k_3^* < k_2$. Constant rates and parameters values: $k_1^* = 10^{-7} s^{-1}$, $k_{-1}^* = 10^{-11} s^{-1}$, $k_2^* = k_{-2}^* = 0$, $k_3^* = 10 s^{-1} M^{-1}$, $k_{-3}^* = 10^{-3} s^{-1} M^{-1}$, $k_1 = 10^{-9} s^{-1}$, $k_{-1} = 10^{-16} s^{-1}$, $k_2 = 10^2 s^{-1} M^{-1}$, $k_{-2} = 10^{-4} s^{-1} M^{-1}$, $k_3 = k_{-3} = 0$, $V = V^* = 10L$ and $a = 0.1L \cdot s^{-1}$ (for $T = 275K$ and $T^* = 413K$). Initial conditions: (a) $[A]_0 = 10^{-6}M$, $[L]_0 = 10^{-11}M$, $[D]_0 = 10^{-11}M$, $[A^*]_0 = 10^{-6}M$, $[L^*]_0 = 10^{-11} + 10^{-21}M$ and $[D^*]_0 = 10^{-11}M$, then $ee_0 = 5 \cdot 10^{-11}$, considerably lower than the calculated stochastic excess, $ee_{st} = 1.9 \cdot 10^{-7}$, (b) $[A]_0 = 10^{-11}M$, $[L]_0 = 5 \cdot 10^{-7}M$, $[D]_0 = 5 \cdot 10^{-7}M$, $[A^*]_0 = 10^{-11}M$, $[L^*]_0 = 5 \cdot 10^{-7} + 10^{-20}M$ and $[D^*]_0 = 5 \cdot 10^{-7}M$, corresponding to an $ee_0 = 10^{-14}M$, in this case the calculated stochastic excess, $ee_{st} = 8.8 \cdot 10^{-10}M$, (c) $[A]_0 = 10^{-10}M$, $[L]_0 = 5 \cdot 10^{-4}M$, $[D]_0 = 5 \cdot 10^{-4}M$, $[A^*]_0 = 10^{-10}M$, $[L^*]_0 = 5 \cdot 10^{-4} + 10^{-5}M$ and $[D^*]_0 = 5 \cdot 10^{-4}M$, that is $ee_0 = 0.0099$, substantially higher than the stochastic excess, $ee_{st} = 2.7 \cdot 10^{-11}$.

In such conditions the system can evolve, for certain reaction and system parameters, towards a chiral stationary state, amplifying a tiny initial enantiomeric excess (even lower than the calculated ee_{st}).

Some analogies of such a cycle with those of the proposals of primordial autocatalytic cycles [18], [19], [20], [21] suggest that the quest for the emergence of chirality in chemical evolution should be jointly studied with that of these premetabolic cycles.

- The LES model is able to lead to SMSB (as a stationary final state), in a two-compartment model if both autocatalytic reactions are spatially separated at different temperatures in different compartments but coupled under the action of a continuous internal flow.

Numerical simulations using reasonable chemical parameters suggest that an adequate scenario for such a SMSB would be that of abyssal hydrothermal vents, by virtue of the typical temperature gradients found there and the role of inorganic solids mediating chemical reactions in an enzyme-like role (see Ref. [15]). The more relevant feature of this LES scenario is its ability to exhibit SMSB at very low reactant concentrations: the effect on the SMSB dependence of the minimal internal flow rate and of the reaction rate inequality $k_{-3}^* > k_{-2}$, and the variability range for SMSB of these parameters increases when the total system concentration C decreases, i.e. the probabilities for SMSB to occur increase in a prebiotic scenario.

Bibliography

- [1] Crusats, J., Hochberg, D., Moyano, A. and Ribó, J. M. *ChemPhysChem* **10**, 2123 (2009).
- [2] Jacques, J., Collet, A., and Wilen, S. H. (John Wiley & Sons, New York, 1981).
- [3] Avetisov, V. and Goldanskii, V. *Proc. Natl. Acad. Sci. USA*. **93**, 11435 (1996).
- [4] Frank, F. *Biochim. Biophys. Acta* **11**, 459 (1953).
- [5] Ribó, J. M. and Hochberg, D. *Phys. Lett. A* **373**, 111 (2008).
- [6] Blackmond, D. G. and Matar, O. K. *J. Phys. Chem. B* **112**, 5098 (2008).
- [7] Cintas, P. and Viedma, C. *Chem. Commun.* **47**, 12786 (2011).
- [8] El-Hachemi, Z., Crusats, J., Ribó, J. M., McBride, J. M. and Veintemillas-Verdaguer, S. *Angew. Chem. Int. Ed. Engl.* **50**, 2359 (2011).
- [9] Noorduyn, W. L., Izumi, T., Millemaggi, A., Leeman, M., Meeke, H., van Enkevort, W. J. P., Kellog, R. M., Kaptein, B., Vlieg, E. and Blackmond, D. G. *J. Am. Chem. Soc.* **130**, 1158 (2008).
- [10] Noorduyn, W. L., Vlieg, E., Kellogg, R. M. and Kaptein, B. *Angew. Chem. Int. Ed. Engl.* **48**, 9600 (2009).
- [11] Viedma, C. *Phys. Rev. Lett.* **94**, 065504 (2005).
- [12] Wattis, J. A. D. *Origins Life Evol B* **41**, 133 (2011).
- [13] Viedma, C., Noorduyn, W. L., Ortiz, J. E., de Torres, T. and Cintas, P. *Chem. Commun.* **47**, 671 (2011).
- [14] **Blanco, Celia**, Crusats, J., El-Hachemi, Z., Moyano, A., Hochberg, D. and Ribó, J. M. Spontaneous emergence of chirality in the limited enantioselectivity model: Autocatalytic cycle driven by an external reagent. *ChemPhysChem* **14**, 2432–2440 (2013).
- [15] Ribó, J. M., Crusats, J., El-Hachemi, Z., Moyano, A., **Blanco, Celia** and Hochberg, D. Spontaneous mirror symmetry breaking in the limited enantioselective autocatalysis model: abyssal hydrothermal vents as scenario for the emergence of chirality in prebiotic chemistry. *Astrobiology* **13**, 132–142 (2013).
- [16] **Blanco, Celia**, Ribó, J. M., Crusats, J., El-Hachemi, Z., Moyano, A. and Hochberg, D. Mirror symmetry breaking with limited enantioselective autocatalysis and temperature gradients: a stability survey. *Phys. Chem. Chem. Phys.* **15**, 1546–1556 (2013).
- [17] Chang, R. *Physical Chemistry for the chemical and biological sciences* (University Science Books, Sausalito, 2000).
- [18] Wächtershäuser, G. *Microbiol. Rev.* **52**, 452 (1988).
- [19] Wächtershäuser, G. *Proc. Natl. Acad. Sci. USA*. **87**, 200 (1988).
- [20] Morowitz, H. J., Kostelnik, J. D., Yang, J. and Cody, G. *Proc. Natl. Acad. Sci. USA*. **97**, 7704–7708 (2000).
- [21] Eschenmoser, A. *Tetrahedron* **63**, 12821–12844 (2007).

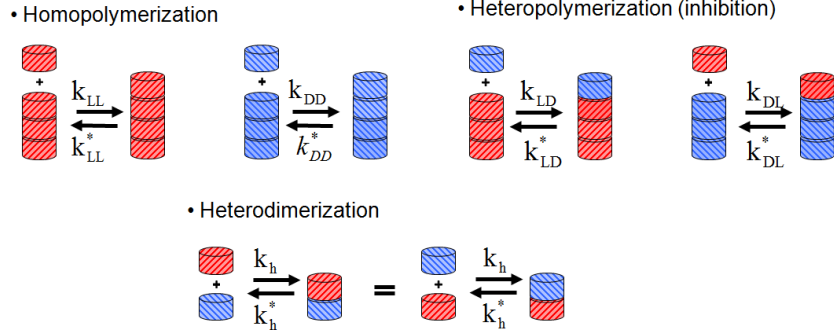
Chapter 4

Chiral polymerization and copolymerization

4.1 Introduction

Biological homochirality of living systems involves large macromolecules, therefore a central point is the relationship of the polymerization process with the emergence of chirality. This hypothesis has inspired recent activity devoted to modeling efforts aimed at understanding mirror symmetry breaking in polymerization of relevance to the origin of life. The models so proposed [1–5] are by and large, elaborate extensions and generalizations of Franks original paradigmatic scheme. [6] Sandars [1] introduced a detailed polymerization process plus the basic elements of enantiomeric cross inhibition as well as a chiral feedback mechanism in which only the largest polymers formed can enhance the production of the monomers from an achiral substrate. He treated basic numerical studies of symmetry breaking and bifurcation properties of this model for various values of the number of repeat units N . All the subsequent models cited here are variations on Sandars' original theme. Brandenburg and coworkers [3] studied the stability and conservation properties of a modified Sandars' model and introduce a reduced $N = 2$ version, including the effects of chiral bias. In ref. [2], they included spatial extent to study the spread and propagation of chiral domains as well as the influence of a background turbulent advection velocity field. The model of Wattis and Coveney [4] differs from Sandars' in that they allow polymers to grow to arbitrary lengths N and the chiral polymers of all lengths, from the dimer and upwards, act catalytically in the breakdown of the achiral source into chiral monomers.

The model introduced by Saito and Hyuga [5] gives rise to homochiral states but differs markedly from Sandars' in that it does not invoke the enantiomeric cross inhibition, allowing instead for reversibility in all the reaction steps. Their model requires open flow, which is the needed element of irreversibility. A different model which stands apart from the above group is that of Plasson et al. [7] They considered a recycled system based on reversible chemical reactions and open only to energy flow and without any (auto)catalytic reactions. A source of constant external energy (the element of irreversibility) is required to activate the monomers. This energy could be introduced into the system in a physical form, say, as high energy photons. A system of this kind, limited to dimerizations, was shown to have nonracemic stable final states for various ranges of the model parameter values and for total concentrations greater than a minimal value. A simple APED model like this (activation-polymerization-epimerization-depolymerization) can be considered for describing the emergence of chiral solutions within a non-catalytic framework for chiral polymerization; the minimal APED model for dimerization can lead to the spontaneous appearance of chiral oscillations. The nature of these oscillations in the enantiomeric excess



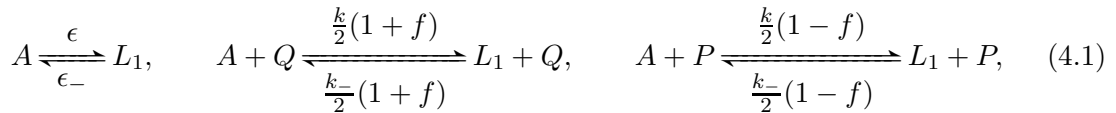
Scheme 4.1: The polymerization model. The L (red) and D (blue) monomers reversibly associate into the growing homochiral polymer chains. The heterodimer formation must be treated separately to avoid double counting and thus ensure that the total system mass is conserved in a closed system.

-as a consequence of oscillations of the concentrations of the associated chemical species-, their relevance and implications in prebiotic chemistry is discussed in Ref. [8].

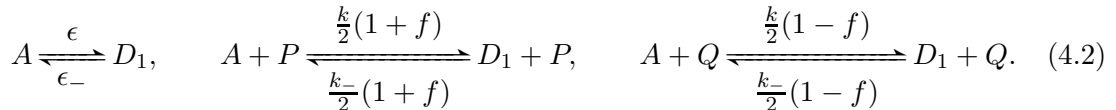
The polymerization models referred to above are defined only for *open flow* systems which exchange matter and energy with the exterior. A constant source of achiral precursor is usually assumed. An unrealistic consequence is that homochiral chains can grow to infinite length. By contrast, most experimental procedures are carried out in *closed and spatially bounded* reaction domains and are initiated in far from equilibrium states [9–16]. It is thus crucial to have models compatible with these experimentally realistic boundary and initial conditions. The most immediate consequences are that polymer chains can grow to a finite maximum length and that the total system mass is constant.

4.2 The polymerization model

The model we introduce and study here, is modified and extended from that of Wattis and Coveney [4] which is in turn, a generalization of Sandar’s scheme [1]. Three salient differences that distinguish our model from these and other previous ones [1–5, 17] are that we (i) consider polymerization in closed systems (so that no matter flow is permitted with an external environment), (ii) we allow for reversible reactions in all the steps and (iii) we also include the formation (and dissociation) of the heterodimer. While heterodimer formation was originally contemplated in ref. [1], it has been silently omitted from all the subsequent models [2–4, 17] that derive therefrom. We assume there is an achiral precursor A which can directly produce the chiral monomers L_1 and D_1 at a slow rate ϵ as well as be consumed in processes in which homopolymers of all lengths catalyze the production of monomers. The specific reaction scheme we study here is composed of the following steps, where $0 \leq f \leq 1$ is the fidelity of the feedback mechanism:

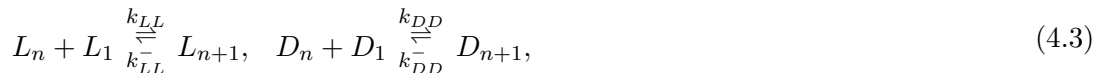


and

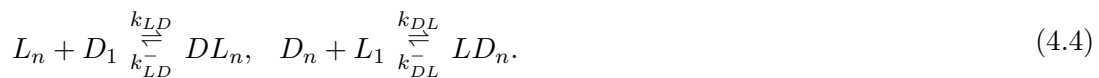


Here $Q = \sum_{n=1}^N nL_n$, and $P = \sum_{n=1}^N nD_n$ represent a measure of the total concentrations of left-handed and right handed polymers. We allow for the monomers themselves to participate in these substrate reactions: hence $N \geq n \geq 1$, where N is the maximum chain length. The central top and bottom reactions in Eqs. (4.1 and 4.2) are enantioselective, whereas those on the right hand side are non-enantioselective: the model contains the features of limited enantioselectivity (see 3).

The following reactions are schematized in Fig.4.1: The monomers combine to form chirally pure polymer chains denoted by L_n and D_n , according to the isodesmic [18] stepwise reactions for $1 \leq n \leq N - 1$,



and inhibition, or the chain end-termination reactions for $N - 1 \geq n \geq 2$:



These upper limits for n specified in Eqs.(4.3,4.4) ensure that the *maximum* length for all oligomers produced (or consumed) by these reaction sets, both the homo- and heterochiral ones, is never greater than N .

In the remainder of the present chapter we will consider here the natural and chiral symmetric rate assignments $k_{LL} = k_{DD}$, $k_{LD} = k_{DL}$ and likewise for the inverse rates, $k_{LL}^- = k_{DD}^-$ and $k_{LD}^- = k_{DL}^-$ ¹. There are then four independent rate constants.

We include a separate reaction for the heterodimer formation and dissociation:



where $k_h = (k_{LD} + k_{DL})/2$ and $k_h^- = (k_{LD}^- + k_{DL}^-)/2$. Note that these latter two rate constants are automatically determined from the above choice and that $L_1D_1 = D_1L_1$ of course. This completes the specification of the model's reactions.

Note that in the elementary reaction steps, in the rate constants, and in the corresponding differential rate equations (see Ref. [19]), the left-right symmetry of the model is manifest, that is, possesses a discrete Z_2 symmetry. This symmetry can be broken spontaneously by the dynamical solutions of the differential rate equations, thus this model is apt for studying spontaneous mirror symmetry breaking.

By lifting the Z_2 degeneracy in the reaction rates, e.g., allowing for $k_{LL} \neq k_{DD}$ and thus leading to more independent rate constants for describing the reaction set, we could study the influence of explicit chiral bias in the model. As this is not the aim of this work, we will not consider it here.

The differences in the Gibbs free energy ΔG^0 between initial and final states should be the same in all the reactions listed in Eqs.(4.1 and 4.2), which implies the thermodynamic constraint on the following forward and reverse reaction constants (see also [5]): $\frac{\epsilon}{\epsilon_-} = \frac{k}{k_-}$.

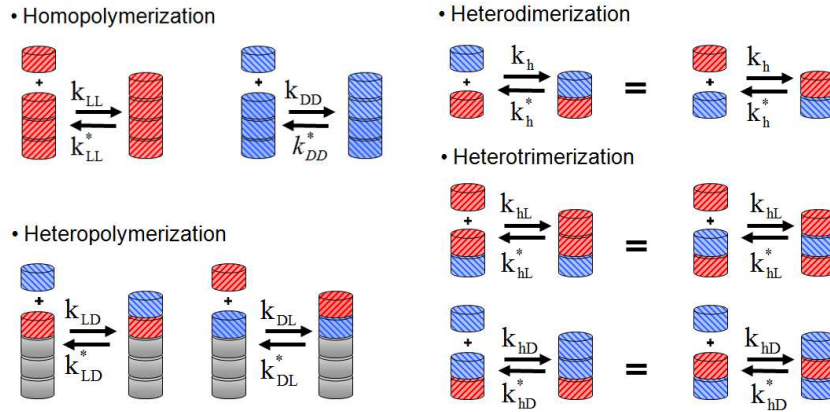
¹Though not considered here, the effects of explicit chiral bias (e.g., that induced by external physical fields) can also be studied with this model by lifting the Z_2 degeneracy in the reaction rates, e.g., allowing for $k_{LL} \neq k_{DD}$, etc., leading to a maximum of eight independent rate constants for describing the reaction set in Eqs. (4.3,4.4).

Rate-equation theory as employed in chemical kinetics is used to describe the differential rate equations of the achiral source, the monomers, as well as the homo- and heterochiral oligomers belonging to this reaction network (the kinetic equations for the concentrations that follow from these elementary steps can be found in Ref. [19]). The system is described by a set of $4N - 2$ kinetic equations that satisfy (the overdot denotes time derivative):

$$[\dot{A}] + 2[\dot{H}] + \sum_{n=1}^N n([\dot{L}_n] + [\dot{D}_n]) + \sum_{n=2}^{N-1} (n+1)([\dot{DL}_n] + [\dot{LD}_n]) = 0. \quad (4.6)$$

4.3 The copolymerization model

Consider now a simple model for the copolymerization of two chemically distinct monomers displaying a wide variety of product sequence compositions. The model we introduce and study here is an appropriately modified and extended version of the one considered a few years ago by Wattis and Coveney [20]. The main important differences



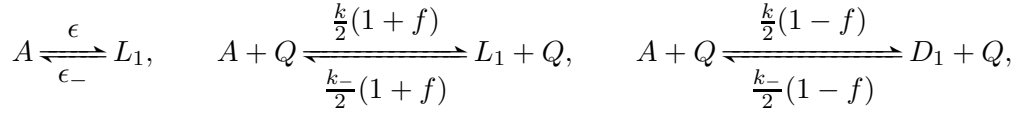
Scheme 4.2: The copolymerization model. The (L)-chiral (red) and (D)-chiral (blue) monomers reversibly associate into the growing homochiral (top) or heterochiral (bottom) copolymer chains. Because the system is *closed*, both the heterodimer (second line) and hetero-trimer (third and fourth lines) reactions must be treated separately to avoid double counting and thus ensure that the total system mass is conserved in a closed system (see text for an explanation).

compared to prior and related models are that we (1) consider polymerization in *closed* systems [21], so that no matter flow is permitted with an external environment— and (2) we allow for reversible monomer association steps. We also correctly include the formation (and dissociation) of the heterodimer [21]. It turns out this must be treated on a separate basis in order to avoid double counting, which if left unchecked, would lead to a violation in the constant mass constraint. Once the heterodimer is treated correctly, this implies that the hetero-trimer must also be treated separately. Beyond this, the remainder of the hetero-oligomers can be treated in a uniform way.

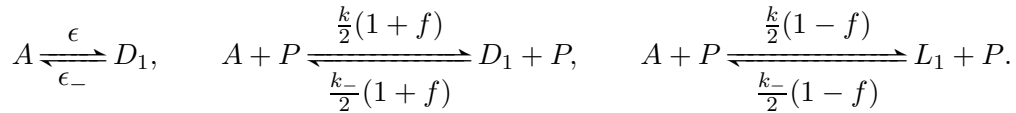
First, we introduce the notation to be used. Polymers are classified by three quantities: the number of L monomers of which it is composed (subscript l), the number of D monomers which it contains (subscript r) and the final or terminal monomer in the chain, denoted by a superscript. In this scheme, the monomers are denoted by $L \equiv C_{1,0}^L$ and $D \equiv C_{0,1}^D$; pure homopolymers are denoted by $L_n \equiv C_{n,0}^L$ and $D_n \equiv C_{0,n}^D$; all copolymer

chains $C_{l,d}^L$ or $C_{l,d}^D$ with $l, d \geq 1$ are heteropolymers. Note also that chains of the form $C_{0,n}^L$ and $C_{n,0}^D$ are forbidden. The corresponding time-dependent concentrations are denoted by lower case variables: e.g., $c_{l,d}^L(t)$ and $c_{l,d}^D(t)$.

As before, we assume there is an achiral precursor A which can directly produce the chiral monomers $C_{1,0}^L$ and $C_{0,1}^D$ as well as be consumed in processes in which homopolymers of all lengths catalyze the production of monomers. :



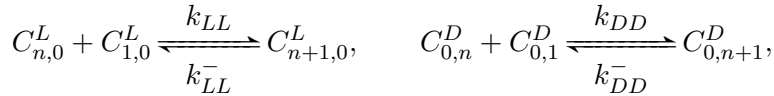
and



In this case, $Q = \sum_{r=1}^N r C_{r,0}^L$, and $P = \sum_{s=1}^N s C_{0,s}^D$. Again, we must obey the thermodynamic constraint $\frac{\epsilon}{\epsilon_-} = \frac{k}{k_-}$.

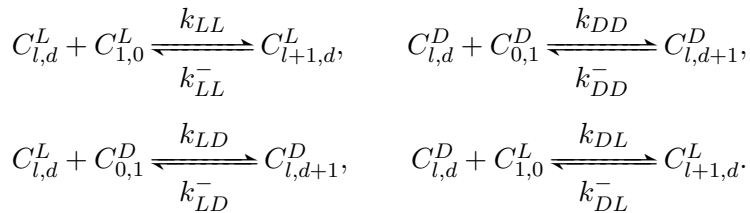
The overall basic scheme must be broken down into several special subcases, especially important so as to avoid undesired double counting of the heterodimer and heterotrimer reactions, see Fig. 4.2.

The formation of chirally pure polymer chains denoted by $c_{n,0}^L$ and $c_{0,n}^D$, for $1 \leq n \leq N-1$ is described by the homo-polymerization reactions:

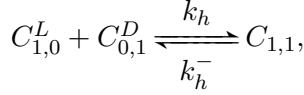


where N is the maximum chain length permitted. In contrast to the polymerization model presented in 4, in the present model, we assume that chains can continue growing by adding monomers of both configurations.

In contrast to the polymerization model presented in 4.2, in this model (as in the original one for open systems [20]), the polymeric chains that have taken up the "wrong" chirality monomer can continue to grow. Thus, we allow for the further growth of these chains by adding monomers of either chirality:

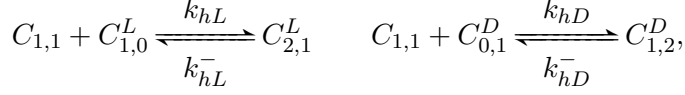


Note that even if we have the information about the composition, we can only know the chirality of the last monomer attached to the chain, we have no information regarding the specific *sequence*. Then, for all practical purposes, $C_{1,1}^L \equiv C_{1,1}^D$, and this suggests using the following notation: $C_{1,1} \equiv C_{1,1}^L \equiv C_{1,1}^D$, and to define a unique direct constant rate: $k_h = \frac{k_{LD} + k_{DL}}{2}$, and an inverse one $k_h^- = \frac{k_{LD}^- + k_{DL}^-}{2}$. Note that if $k_{LD} = k_{DL}$, then $k_h = k_{LD} = k_{DL}$. Due to these characteristics, we will treat the heterodimer in a different way compared with the other hetero-polymers. The reaction of the heterodimer formation is therefore:



implying the thermodynamic constraint $\frac{k_{LD}}{k_{LD}^-} = \frac{k_{DL}}{k_{DL}^-}$.

The heteropolymers formed from the addition of a monomer to a heterodimer:



where $k_{hL} = \frac{k_{LL} + k_{DL}}{2}$, $k_{hD} = \frac{k_{DD} + k_{LD}}{2}$ and $k_{hL}^- = \frac{k_{LL}^- + k_{DL}^-}{2}$, $k_{hD}^- = \frac{k_{DD}^- + k_{LD}^-}{2}$. Note that if $k_{LL} = k_{DD}$ and $k_{LD} = k_{DL}$, then $k_{hL} = k_{hD}$, and if $k_{LL}^- = k_{DD}^-$ and $k_{LD}^- = k_{DL}^-$, then $k_{hL}^- = k_{hD}^-$. Again, the left-right symmetry of the model is manifest. It possesses a discrete Z_2 symmetry that can be broken spontaneously by the dynamical solutions of the differential rate equations.

The total number of equations for describing the system as a function of maximum chain length N is:

$$\#eqs = 6 + 2(N-1) + 2(N-2) + 2(N-3) + (N^2 - 7N + 12) + 2(N-3) = N(N+1),$$

so from the computational point of view, the number of equations grows quadratically with the maximum chain length N . As remarked earlier, the complete reaction scheme must satisfy mass conservation in a closed system, implying that the mass variation rate must be strictly zero:

$$\begin{aligned} 0 &= 2\dot{c}_{1,1} + 3(\dot{c}_{2,1}^L + \dot{c}_{1,2}^D) + \sum_{n=1}^N n(\dot{c}_{n,0}^L + \dot{c}_{0,n}^D) + \sum_{n=2}^{N-1} (n+1)(\dot{c}_{1,n}^L + \dot{c}_{n,1}^D) \\ &+ \sum_{n=2}^{N-2} (n+2)(\dot{c}_{2,n}^L + \dot{c}_{n,2}^D) + \sum_{l=3}^{N-1} \sum_{d=1}^{N-1} (l+d)(\dot{c}_{l,d}^L + \dot{c}_{l,d}^D), \end{aligned} \quad (4.7)$$

where the overdot stands for the time-derivative.

4.4 Results

Results are quantified in terms of a variety of convenient chiral measures. The percent enantiomeric excess values of the oligomers with homochiral sequence are calculated according to ($2 \leq n \leq N$)

$$ee_n = \frac{[L_n] - [D_n]}{[L_n] + [D_n]}. \quad (4.8)$$

A global or ensemble-averaged measure of the degree of symmetry breaking is provided by the number-weighted enantiomeric excess η :

$$\eta = \frac{\sum_{n=2}^N n([L_n] - [D_n])}{\sum_{n=2}^N n([L_n] + [D_n])}. \quad (4.9)$$

The average chain length of the homopolymers is given by:

$$\bar{n} = \frac{\sum_{n=2}^N n([L_n] + [D_n])}{\sum_{n=2}^N ([L_n] + [D_n])}, \quad (4.10)$$

and the root mean square deviation in the homochiral chain length are:

$$\overline{(n^2)}^{1/2} \equiv \sqrt{\langle (n - \bar{n})^2 \rangle} = \sqrt{\langle n^2 \rangle - \langle n \rangle^2}. \quad (4.11)$$

The angular brackets $\langle \rangle$ denote averaging with respect to the ensemble $\sum_n ([L_n] + [D_n])$, similar to Eq.(4.10). It is important to remember that these are all time-dependent quantities.

The entropy production rate in an irreversible process is a measure of the dissipation in that process. At equilibrium, the entropy production rate vanishes and is an extremum [22]. This production has been investigated recently for reversible versions of the Frank model [23, 24]. In those simple models, the behavior of the entropy produced near the chiral symmetry breaking transition as well as its subsequent temporal development, depends on whether the chemical system is open or closed to matter flow (see Ref. [19]).

For reactions obeying mass action kinetics, the entropy produced in any chemical reaction can be calculated straightforwardly in terms of the individual elementary reaction rates [22, 25]. The rate of entropy production is the sum over the difference of the forward (R_{jf}) and reverse (R_{jr}) reaction rates multiplied by the natural logarithm of the ratio of the forward and reverse rates [25]

$$\sigma(t) = R^* \sum_j (R_{jf} - R_{jr}) \ln \left(\frac{R_{jf}}{R_{jr}} \right) \geq 0, \quad (4.12)$$

where the sum runs over each elementary reaction step j , and $R^* = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$ is the gas constant. Since our reaction scheme is set up for closed systems, equilibrium is reached after a racemization time scale t_{racem} is reached, which suggests a further measure is provided by the total net entropy produced from the start of the reactions through chiral symmetry breaking, then on to final racemization, when the system reaches chemical and thermodynamic equilibrium and $\sigma(t > t_{racem}) = 0$

$$\sigma_T = \int_0^{t_{racem}} \sigma(u) du. \quad (4.13)$$

This quantifies the total dissipation over the complete history of the chemical transformations under study.

The central point here is studying the sensitivity of the above reversible polymerization and copolymerization networks to minuscule initial enantiomeric excesses, so a very dilute initial concentration of a non racemic mixture was employed in the calculations. Results corresponding to the polymerization model (section 4.2) are shown in Fig. 4.1a. The enantiomeric excess η (solid black lines), starts off at zero value until a time on the order of $t \sim 10$ at which the excess increases rapidly to nearly 100% at SMSB. This is followed by a gradual stepwise decrease characterized by the appearance of quasi-plateaus of approximate constancy: η falls to about 90% at $t \sim 10^3$, then to about 60% at $t \sim 10^6$, staying approximately level until the final decrease to zero occurring at a time on the order of $t \sim 10^{11} - 10^{12}$. The system has racemized on this time scale. Results for the copolymerization model (section 4.3) are shown in Fig. 4.1b. In this case, the behaviour of the enantiomeric excess η (start at zero, rapid increase to near 100%, stepwise decrease to zero, final racemization) is qualitatively similar to that shown in Fig.4.1a. No appreciable differences in η can be discerned when we include the monomer, that is, start the sums at $n = 1$ (slow chiral erosion proceeding through quasi-steady plateaus).

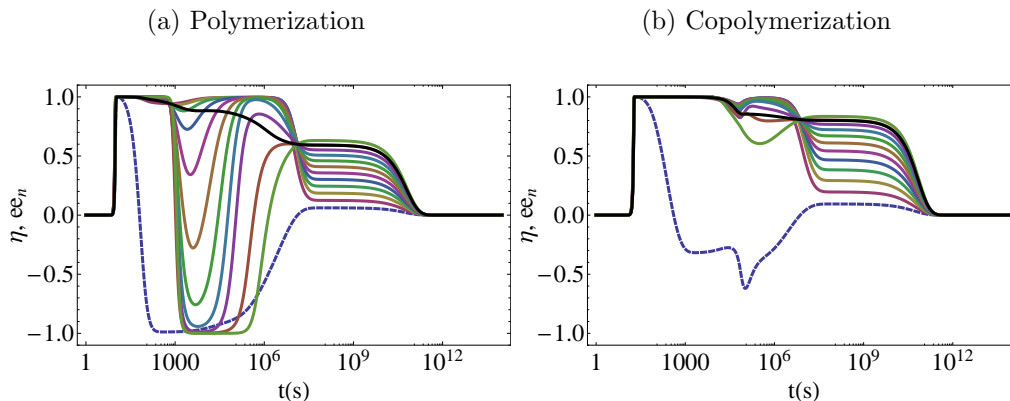


Figure 4.1: Time evolution of the average enantiomeric excess η (black lines) and the individual chain-length dependent enantiomeric excesses ee_n (family of solid curves). The dashed curve shows the chiral excess for the monomers ee_1 . Initial concentrations: $[L_1]_0 = 10^{-6} + 10^{-15}M$ and $[D_1]_0 = 10^{-6}M$ ($ee_0 = 5 \cdot 10^{-10}$, $ee_{st} = 6.2 \cdot 10^{-10}$), $[A]_0 = 2M$, the rest of initial concentrations are zero. Reaction rates: $\epsilon = 2 \cdot 10^{-5}s^{-1}$, $\epsilon_- = 10^{-10}s^{-1}$, $k = 2.0s^{-1}M^{-1}$, $k_- = 10^{-5}s^{-1}M^{-1}$, $f = 0.9$, $k_{LL} = k_{DD} = 1.0s^{-1}M^{-1}$, $k_{LD} = k_{DL} = 10^3s^{-1}M^{-1}$, $k_{LD}^- = k_{DL}^- = 1.0s^{-1}$, $k_{LL}^- = k_{DD}^- = 10^{-5}s^{-1}$. For (a) near $t \simeq 10^4$, the sequence of curves corresponds to $n = 2$ to $n = 12$ top to bottom and for (b) near $t = 10^9$ from bottom to top.

Since the enantiomeric cross inhibition $k_{LD} = k_{DL}$ is a determining factor in this model, by way of contrast, we consider a second $N = 12$ run with a much lower mutual inhibition than employed above, namely $k_{LD} = k_{DL} = 20s^{-1}M^{-1}$, and with the following inverse rates all set equal $k_{LL}^- = k_{DD}^- = k_{LD}^- = k_{DL}^- = 10^{-6}s^{-1}$, but keeping the remainder of the rates as before and with the same initial concentrations and excess. In this situation, the symmetry breaking occurs at a later time (see Ref. [19])

The individual percent chain-length dependent enantiomeric excesses, Eq.(4.8) plotted in Fig. 4.1 (coloured lines) reveal a remarkable complex dynamic behavior: the individual ee 's follow a common curve from initialization to chiral symmetry breaking and remain together at nearly 100% until the common curve begins to split up into its constituents. Then, the percent chiral excess of each length homochiral chain behaves differently, until they again coalesce into a single curve upon final racemization.

Only in Fig.4.1a the largest length chains pass from *positive* then to *negative* values of the excess (the $N = 12, 11, 10$ chains exhibit nearly -100% excess during the period from 10^3 to 10^4 and beyond): there has been a chiral *sign reversal* in the excess corresponding to the largest chains. This holds also for the monomer ee_1 , plotted in the dashed curve. Except for the monomer, the individual excesses then all increase back to positive values at $t \sim 10^6$, then from $t \sim 10^7$ to $t \sim 10^{11}$, the excess increases sequentially as a function of the chain length n until racemization, where they all collapse to zero. The temporal behavior of the enantiomeric excesses of the largest chains $n = 12, 11, 10, 9, 8, 7$ is reminiscent of strongly *damped oscillations*. In Fig. 4.1b, the individual ee 's though behaves qualitatively similar to those obtained in Fig.4.1a, but there is a remarkable difference: except for the monomers, the individual excesses have all them positive values at all times.

Static "snapshots" of this dynamic behavior complement the evolution of the chain length dependent enantiomeric excesses. In Fig. 4.2 we display the enantiomeric excess versus the number of chiral repeat units at selected time slices. In Fig.4.2a, in the left-hand graph, the ee 's are all at 100% for all the chains (contrast with Fig.4.1a). The next

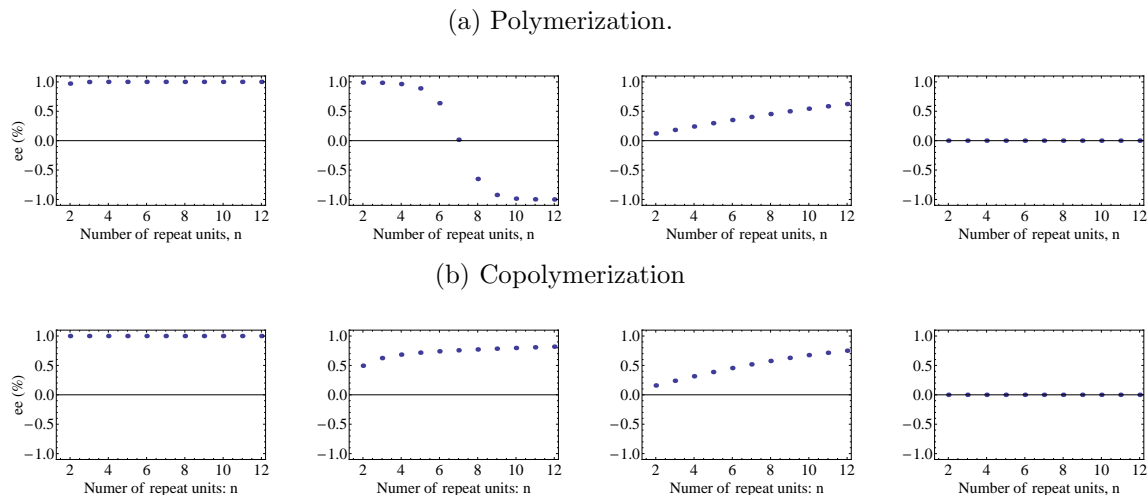


Figure 4.2: Different time slices or "snap-shots" of Fig.4.1a and Fig.4.1b showing dependence of the chain-length dependent enantiomeric excesses ee_n for $n \geq 2$, at different selected time scales for (a) $t = 100$, $t = 10^4$, $t = 10^9$, $t = 10^{13}$ and (b) $t = 100$, $t = 10^7$, $t = 10^{10}$, $t = 10^{13}$

graph, corresponding to $t = 10^4$, shows the sign reversal tendency as a function of chain length, with the full reversal (-100%) being attained for the largest homochiral chains. The following graph, corresponding to $t = 10^9$ shows the monotone increase of ee with chain length. Finally, the righthand most graph shows that racemization has set in by $t = 10^{13}$.

The qualitative behaviors depicted in the second and third snapshot have been reported in two recent and independent polymerization experiments [15] [12] (see Ref. [19]). The tendency of the sign reversal in ee (from positive to negative values) as a function of chain length has been observed in the polymerization of racemic valine (Val-NCA) and leucine (Leu-NCA) in water subject to chiral initiators [15]. By contrast, the monotonic increase of the percent ee with chain length has been measured in independent chiral amplification experiments using leucine and glycine in water [12] starting with a 20% initial enantiomeric excess of the L_1 monomer.

In Fig.4.2b, in the left-hand graph, the ee 's are all at 100% for all the chains at $t = 100$ (contrast with Fig.4.1b). The two next graphs, corresponding to $t = 10^7$ and $t = 10^{10}$, shows the decreasing tendency as a function of chain length, the largest homochiral chains remain at values near 100% longer. Finally, the righthand most graph shows that racemization has set in by $t = 10^{13}$.

Additional information regarding the homo-oligomer composition of the chemical system is provided by the average homochiral chain length $\langle n \rangle$, see Eq.(4.10). We plot this in Fig. 4.3 along with the standard deviation about the mean, Eq.(4.11). The final stable values of the mean and RMS values are $\bar{n} = 11.0 \pm 1.4$ for the polymerization model and $\bar{n} = 10.9 \pm 1.5$, for the copolymerization model. This demonstrates that the final racemic mixture is dominated by the longer length homopolymer chains, and this is the final equilibrium configuration.

The behavior of entropy in polymerization models is rarely discussed, [26] and has not been addressed previously for mirror symmetry breaking in chiral polymerization. The entropy produced in a chemical reaction initiated out of equilibrium gives a measure of

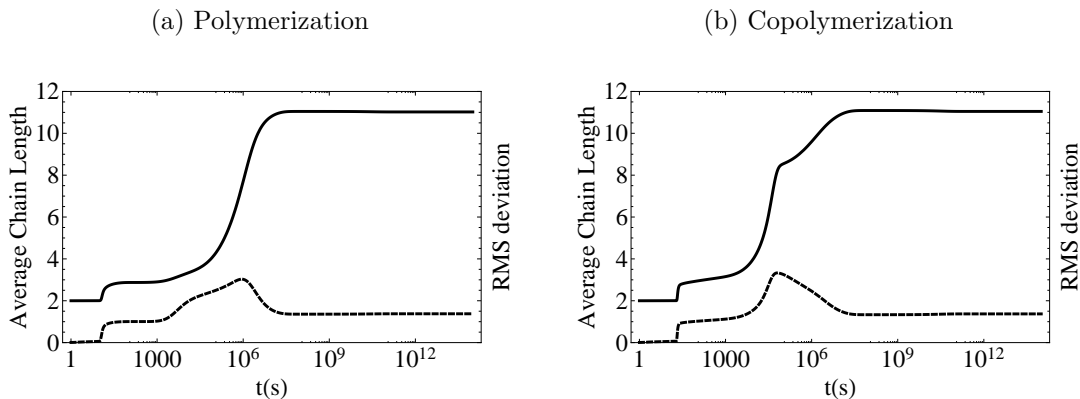


Figure 4.3: The time evolution (logarithmic scale) of the average or mean homopolymer chain length \bar{n} (upper curve) and the corresponding root-mean-square deviation $\langle \overline{n^2} \rangle^{\frac{1}{2}}$ from the mean value (lower dashed curve).

the dissipation during the approach to final equilibrium. In Ref. [19] the rate of entropy production in chiral polymerization is calculated following Eq.4.12 when the system undergoes a chiral phase transition as well at the later stages when the system reaches final chemical equilibrium upon racemization.

The sum in Eq.(4.12) contains $2N + 4$ terms for both models; so it is possible to determine which specific steps of the full reaction network provide the leading contributions to σ by representing them individually.

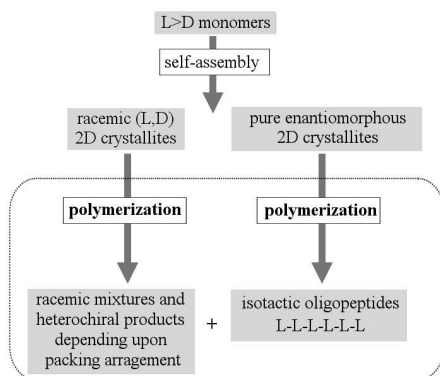
For both, the polymerization and copolymerization, using the same initial concentrations and reaction rates, the leading contribution to the entropy production comes from the monomer catalysis steps eqn (1), followed by the polymerization itself eqn (4), next by the mutual inhibition reactions: heterodimer formation eqn (6) and the "end-spoiled" cross-inhibition reactions eqn (5). The least important contribution comes from the direct production of monomers from the achiral substrate.

For a lower heterochiral inhibition, the entropy production peaks well before the mirror symmetry is broken. The catalysis still yields the major contribution to this peak, but the second and third most important contributions are now formation of end-chain spoiled oligomers followed by the polymerization, exactly opposite to the previous run employing the much higher mutual inhibition. The peak in σ is due principally to monomer catalysis, and not symmetry breaking.

Interpretation of experimental data

Polymerization reactions of racemic mixtures of monomers in solution are typically expected to yield polymers composed of random sequences of the left- and right-handed repeat units following a binomial or Bernoulli distribution (see also Fig.1.1). Thus the probability for obtaining oligomers with homochiral sequence becomes negligible with increasing length [27]. In contrast, Luisi and coworkers [10, 12, 28, 29] have reported the polymerization of racemic α -amino acids in solution which yields small amounts of oligopeptides of homochiral sequence whose abundances with respect to the heterochiral chains exhibit a slight departure from the binomial distribution. This problem of the random distribution can be overcome by catalyzed polymerization of amphiphilic amino acids, in racemic and

nonracemic forms, which self-assemble into two-dimensional ordered crystallites at the air-water interface [30]. Based on a process involving self-assembly followed by lattice controlled polymerization, Lahav and coworkers recently proposed a general scenario for the generation of homochiral oligopeptides of a single handedness from non-racemic mixtures of activated alpha amino acids [30].



Scheme 4.3

We are interested in the lattice-controlled polymerization reactions proposed by those authors. The proposed experimental scheme, represented in Fig.4.3 (Fig.15 in Ref. [31]) starts from an initial excess, say $L > D$ of monomers which undergoes an initial self-assembly process into two types of two-dimensional crystallites at the air/water interface. Once formed, each one of these two crystal phases participates in the control of a subsequent type of polymerization. Thus, the racemic crystallites polymerize racemic mixtures of oligomers and the heterochiral products, whereas the other pure enantiomorphous crystallite controls the polymerization of the isotactic chains, these are formed from the monomer in excess (L , in this exam-

ple). Our attempt to fit the experimental data to the copolymerization model² is concerned with the effective polymerization controlled by each of the two crystalline phases. Different chemical model systems were used in the experiments: namely γ -stearyl-glutamic thioethyl ester ($C_{18} - TE - Glu$), N^ϵ -stearyl-lysine thioethyl ester ($C_{18} - TE - Lys$), γ -stearyl-glutamic acid N-carboxyanhydride ($C_{18} - Glu - NCA$) and γ -stearyl-glutamic thioacid ($C_{18} - thio - Glu$), varying both their initial compositions and for various choices of catalyst.

Employing the same initial concentrations as in the experiments, and after setting all the inverse reaction rates to a unique value, $k_{LL}^- = k_{DD}^- = k_{LD}^- = k_{DL}^- = 10^{-10} s^{-1}$, and the cross inhibition rates equal to unity ($k_{LD} = k_{DL} = 1 s^{-1} M^{-1}$), the next step is to search for the reaction rates $k_{LL} = k_{DD}$ leading to the best fits. Results from fitting the model to the data indicate that the maximum chain length N does not play a significant role, the Pearson product-moment correlation coefficient, r , remains the same for $N = 12, 14, 16, 18, 20$, so we set $N = 12$ for all compounds³.

In one set of experiments, the authors reported MALDI-TOF analysis of the oligopeptides formed at the air-water interface from racemic mixtures $L : D = 1 : 1$ of the monomers for the various model systems and catalysts. The best correlation data for the racemic C_{18} -TE-Glu system, with the I_2/KI catalyst are found for $k_{LL} = k_{DD} = 1.7(s^{-1} M^{-1})$ (for the time scale $t = 10^{11}(s)$). Exactly by the same process, the best correlation data for the racemic C_{18} -TE-Lys are found for $k_{LL} = k_{DD} = 2.3(s^{-1} M^{-1})$ and for $k_{LL} = k_{DD} = 1.3(s^{-1} M^{-1})$ when adding I_2/KI and $AgNO_3$ as catalyst ($t = 10^{10}(s)$ and $t = 10^{11}(s)$), respectively. Finally, we fit our copolymerization model to the C_{18} -thio-Glu

²The model may be termed *effective* in the following sense: it presupposes or takes as given the *prior* formation of the self-assembled 2D crystallites at the air-water interface and is concerned exclusively with the subsequent polymerization reactions.

³Since the number of independent equations scales as N^2 , this represents an important reduction on computer time and the memory used.

	Copolymerization model							Binomial	
	Fits for each subgroup (n)						Global fit	Global fit	
	di	tri	tetra	penta	hexa	hepta			
<i>C</i> ₁₈ - TE-Glu (L:D) 1:1	0.92	0.96	0.80	0.84	-	-	0.93	0.75	
	(L:D) 4:6	0.86	0.89	0.93	0.99	-	-	0.94	0.75
	(L:D) 3:7	0.95	0.94	0.96	0.99	0.99	-	0.95	0.85
<i>C</i> ₁₈ - TE-Lys (L:D) 1:1	0.98	1	0.03	0.88	0.76	0.84	-	0.8	
	(L:D) 4:6	0.78	1	0.87	0.90	0.84	0.97	0.89	0.46
	(L:D) 3:7	0.93	1	0.95	0.97	0.99	-	0.94	0.65
<i>C</i> ₁₈ - thio-Glu (L:D) 1:1	1	1	1	0.98	0.97	0.95	-	0.98	
	(L:D) 4:6	0.93	0.98	0.93	0.92	0.92	0.91	0.91	0.93
	(L:D) 3:7	0.89	1	0.99	0.99	0.98	-	0.96	0.97

Table 4.1: Comparative fits between the copolymerization model and the binomial distribution to the experimental relative abundances (Pearson product-moment correlation coefficient, r) measured for racemic and non-racemic mixtures of *C*₁₈-*TE-Glu*, *C*₁₈-*TE-Lys* and *C*₁₈-*thio-Glu*

experimental relative abundances. The authors of the experiments affirmed that this compound undergoes a truly random polymerization [30], so fits from our model are expected to be slightly less satisfactory than those for the binomial distribution function. Setting the inverse reaction rates and the cross inhibition as indicated above, then the best correlation coefficients are found for $k_{LL} = k_{DD} = 0.4(s^{-1}M^{-1})$. The instant or time-scale leading to these numerical values is $t = 10^{10}(s)$. The resulting data correlations are shown in Table 4.1, the latter gives a detailed comparison of the best fits between individual subfamilies and the overall global fit. In a second set of experiments, the authors reported MALDI-TOF analysis of the oligopeptides formed at the air-water interface from non-racemic mixtures of the monomers for the same model systems. No catalysts were employed there. The best correlations factors for both chirally enriched mixture cases (20% and 40% excesses) in the case of the *C*₁₈-*TE-Glu* system are found for the same rates, that is for $k_{LL} = k_{DD} = 2(s^{-1}M^{-1})$ (at $t = 10^{11}(s)$). For the chiral mixtures of *C*₁₈-*TE-Lys* we found the best fits for the dynamics corresponding to $k_{LL} = k_{DD} = 2.5(s^{-1}M^{-1})$ (at $t = 10^{10}(s)$). In the case of nonracemic *C*₁₈-*thio-Glu*, the best correlation coefficients are found for the same values of the reaction rates that we found in the racemic case, namely for $k_{LL} = k_{DD} = 0.4(s^{-1}M^{-1})$. As to be expected and as shown there, the correlation factors for the global fit to the binomial distribution function are slighter better than those for any simulation we could perform with the copolymerization model, so we reconfirm what was claimed by the authors of the experimental work: namely that the *C*₁₈-*thio-Glu* system polymerizes randomly. The results for these values are shown in Table 4.1. Relative abundances and enantiomeric excesses obtained for 4:6 and 3:7 (L:D) mixtures in the four cases can be found in Ref. [19].

4.5 Conclusions

- A strong but temporary chiral amplification can take place in a reversible model of chiral polymerization closed to matter flow and subject to constraints imposed by microreversibility. Strong mutual inhibition is required to amplify the initial ee to large values, very similar to what we found for the reversible Frank model in closed systems (see Chapter 2).

- The chain length dependent enantiomeric excesses is clearly a time dependent phenomena, tending to exhibit a damped oscillatory behavior before the onset of final racemization.⁴ The implications for chirality transmission are far reaching: "memory" of the sign of the initial chiral fluctuation is washed-out by the oscillations in the enantiomeric excess, adding another heretofore unexpected element of randomness to the process. These oscillations cease as the system approaches its equilibrium state.

- Computation of the average homopolymer length indicates that the final racemic state is dominated by the longest available chains (see Fig.4.3).

- The rate of entropy production per unit volume exhibits a peak value either before or near the vicinity of the chiral symmetry breaking transition. The leading contribution to the entropy production comes from the monomer catalysis steps, followed by the polymerization itself, next by the mutual inhibition reactions: heterodimer formation and the "end-spoiled" cross-inhibition reactions. The least important contribution comes from the direct production of monomers from the achiral substrate.

- The copolymerization model is defined for fully reversible reactions and this implies that some of the reaction rates must obey a corresponding constraint as dictated by microreversibility. The model is appropriate for closed systems under thermodynamic control.

- The forward rates of adding the same chirality monomer to the end of the growing chain are found to be greater than those for addition of a wrong chirality monomer (i.e. $k_{LL} = k_{DD} > k_{LD} = k_{DL} = 1$), except for the model system $C_{18} - thio - Glu$, serving as reference for random polymerization.

- The Pearson product-moment correlation coefficient r between experimental and numerical data is greater for the copolymerization model than for the binomial distribution, except for the $C_{18} - thio - Glu$, which truly polymerizes randomly (see Table 4.1).

- A very positive feature of the copolymerization model is the robustness of the fits with respect to differing initial imbalances of the enantiomers. The correlation between calculated and experimental relative abundances is greater for the initially non-racemic situations (the higher the initially chiral enrichment of the mixture is, the better the copolymerization model reproduces the chemical data). The results obtained here lead us to affirm that the model systems considered all undergo a non-random polymerization, as was asserted by the authors of the experiments [30].

- The model also qualitatively reproduces the behavior of the enantiomeric excess ee , its increase with the length of the chains and the enhancement of the ee of the corresponding initial mixture of monomers (see Ref. [21]). In conclusion then, we may therefore

⁴Oscillatory dynamics in chemical reactions has been observed experimentally, and analyzed theoretically and numerically in simple model systems [32]; as far as we are aware, this behavior has not been revealed previously as a valid dynamical solution in polymerization models.

assert that our simple copolymerization scheme does provide an accurate course-grained description of the lattice-controlled polymerization reported in Ref [30].

Bibliography

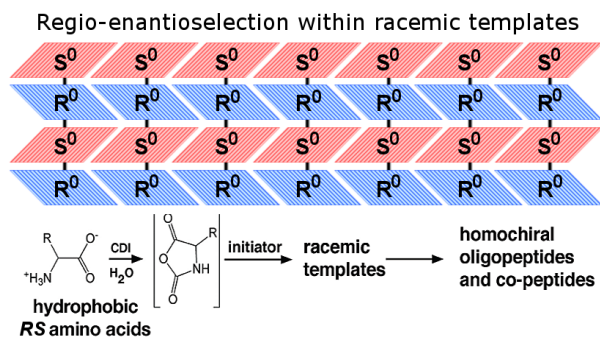
- [1] Sandars, P. *Origins Life Evol B* **33**, 575 (2003).
- [2] Brandenburg, A. and Multamäki, T. *Int. J. Astrobiol.* **3**, 209 (2004).
- [3] Brandenburg, A., Andersen, A., Hofner, S. and Nilsson, M. *Origins Life Evol B* **35**, 225 (2005).
- [4] Wattis, J. and Coveney, P. **35**, 243 (2005).
- [5] Saito, Y. and Hyuga, H. *J. Phys. Soc. JPN* **74**, 1629 (2005).
- [6] Frank, F. *Biochim. Biophys. Acta* **11**, 459 (1953).
- [7] Plasson, R., Bersini, H. and Commeyras, A. *P. Natl. Acad. Sci. USA.* **101**, 16733 (2004).
- [8] Stich, M., **Blanco, Celia** and Hochberg, D. Chiral and chemical oscillations in a simple dimerization model. *Phys. Chem. Chem. Phys.* **15**, 255–261 (2013).
- [9] Illos, R. A., Bisogno, F. R., Clodic, G., Bolbach, G., Weissbuch, I. and Lahav, M. *J. Am. Chem. Soc.* **130**, 8651 (2008).
- [10] Blocher, M., Hitz, T. and Luisi, P. *Helv. Chim. Acta* **84**, 842 (2001).
- [11] Nery, J., Bolbach, G., Weissbuch, I. and Lahav, M. *Angew. Chem. Int. Edit.* **42**, 2157 (2003).
- [12] Hitz, T. and Luisi, P. *Helv. Chim. Acta* **86**, 1423 (2003).
- [13] Hitz, T. and Luisi, P. *Origins Life Evol B* **34**, 93 (2004).
- [14] Rubinstein, I., Eliash, R., Bolbach, G., Weissbuch, I. and Lahav, M. *Angew. Chem. Int. Edit.* **46**, 3710 (2007).
- [15] Rubinstein, I., Clodic, G., Bolbach, G., Weissbuch, I. and Lahav, M. *Chem-Eur. J.* **14**, 10999 (2008).
- [16] Weissbuch, I., Illos, R. A., Bolbach, G. and Lahav, M. *Accounts Chem. Res.* **42**, 1128 (2009).
- [17] Gleiser, M. and Thorarinson, J. *Origins Life Evol B* **36**, 501 (2006).
Gleiser, M. *Origins Life Evol B* **37**, 235 (2007).
Gleiser, M. and Walker, S. I. *Origins Life Evol B* **38**, 293 (2008).
Gleiser, M., Thorarinson, J. and Walker, S. I. *Origins Life Evol B* **38**, 499 (2008).
- [18] De Greef, T. F. A., Smulders, M. M. J., Wolffs, M., Schenning, A. P. H. J., Sijbesma, R. P. and Meijer, E. W. *Chem. Rev.* **109**, 5687 (2009).
- [19] **Blanco, Celia** and Hochberg, D. Chiral polymerization: symmetry breaking and entropy production in closed systems. *Phys. Chem. Chem. Phys.* **13**, 839–849 (2011).
- [20] Wattis, J. A. D. and Coveney, P. V. *J. Phys. Chem. B* **111**, 9546 (2007).
- [21] **Blanco, Celia** and Hochberg, D. Homochiral oligopeptides by chiral amplification: interpretation of experimental data with a copolymerization model. *Phys. Chem. Chem. Phys.* **14**, 2301–2311 (2012).
- [22] Ross, J. and Vlad, M. *J. Phys. Chem. A* **109**, 10607 (2005).
- [23] Kondepudi, D. and Kapcha, L. *Chirality* **20**, 524 (2008).
- [24] Mauksch, M. and Tsogoeva, S. B. *Chem. Phys. Phys. Chem.* **9**, 2359 (2008).
- [25] Kondepudi, D. and Prigogine, I. (New York Wiley, 1998).
- [26] Stier, U. *J. Chem. Phys.* **124**, 024901 (2006).
- [27] Joyce, G., Visser, G., VanBoeckel, C., VanBoom, J., Orgel, L. and Vanwestrenen, J. *Nature* **310**, 602–604 (1984).
- [28] Hitz, T., Blocher, M., Walde, P. and Luisi, P. L. *Macromolecules* **34**, 2443 (2001).
- [29] Hitz, T. and Luisi, P. *Helv. Chim. Acta* **85**, 3975 (2002).
- [30] Zepik, H., Shavit, E., Tang, M., Jensen, T., Kjaer, K., Bolbach, G., Leiserowitz, L., Weissbuch, I. and Lahav, M. *Science* **295**, 1266 (2002).
Weissbuch, I., Zepik, H., Bolbach, G., Shavit, E., Tang, M., Jensen, T., Kjaer, K., Leiserowitz, L. and Lahav, M. *Chem-Eur. J.* **9**, 1782 (2003).
- [31] **Blanco, Celia** and Hochberg, D. Stochastic mirror symmetry breaking: Theoretical models and simulation of experiments **333**, 157–211 (2013).
- [32] Cross, M. and Hohenberg, P. *Rev. Mod. Phys.* **65**, 851 (1993).

Chapter 5

Mirror symmetry breaking via β -sheet-controlled copolymerization: (i) Mass Balance and (ii) Probabilistic Treatment

5.1 Introduction

One scenario for the transition from prebiotic racemic chemistry to chiral biology suggests that homochiral peptides or amino acid chains must have appeared before the onset of the primeval enzymes [1–5]. However, except for a couple of known cases [6, 7], the polymerization of racemic mixtures (i.e., in 1:1 proportions) of monomers in ideal solutions typically yields chains composed of random sequences of both the left and right handed repeat units following a binomial distribution [8].



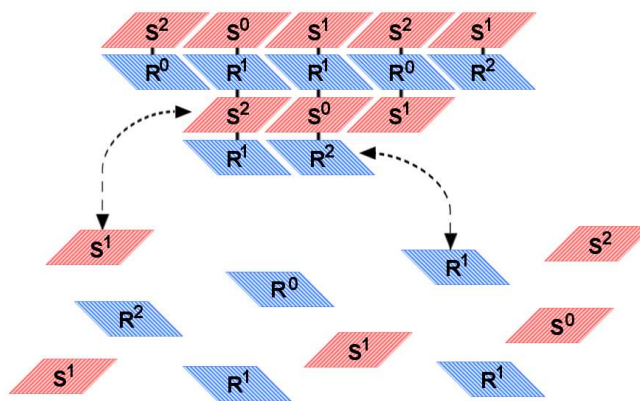
Scheme 5.1: Self assembly of oligopeptides into racemic β -sheets, for the case of a single species (R_0, S_0) of amino acid supplied in ideally racemic proportions. For a full experimental account, see Weissbuch et al. [9].

This *statistical* problem has been overcome recently by the experimental demonstration of the generation of amphiphilic peptides of homochiral sequence, that is, of a single chirality, from racemic compositions or racemates. This route consists of two steps: (1) the formation of racemic parallel or anti-parallel β -sheets either in aqueous solution or in 3-D crystals [9] during the polymerization of racemic hydrophobic α -amino acids (Fig. 5.1) followed by (2) an enantioselective controlled polymerization reaction [10–16]. This process leads to racemic or mirror-symmetric mixtures of isotactic oligopeptides where the chains are composed from amino acid residues of a single handedness (see Fig. 5.1). Fur-

thermore, when racemic mixtures of different types of amino acids were polymerized, isotactic co-peptides of homochiral sequence were generated. The guest (S) and (R) molecules are enantioselectively incorporated into the chains of the (S) and (R) peptides, respectively, however the guest molecules are randomly distributed within the corresponding homochiral chains, see Fig. 5.2.

As a combined result of these two steps, the *sequence* of pairs of co-peptide S and R chains within the growing template will differ from each other, see Fig. 5.2, resulting in non-racemic mixtures of co-peptide polymer chains of different sequences¹. Consequently, by considering the sequences of the peptide chains, a statistical departure from the racemic composition of the library of the peptide chains is created which varies with chain length and with the relative concentrations of the monomers used in the polymerization [11, 12]. This can be appreciated comparing Fig. 5.1 and Fig. 5.2: in the former the β -sheet is globally racemic (no guest amino acids) whereas the latter template is not by virtue of the randomness of the specific amino acid sequences within each homochiral strand, due to the presence of guest amino acids.

In this chapter, we are concerned about the theoretical investigation of multi-component copolymerization controlled by such templates, with this purpose in mind, the two different models here introduced presuppose or take as given the prior formation of the initial templates or β -sheets and is concerned exclusively with the subsequent enantioselective polymerization reactions (the nonlinear template control is implicit throughout our discussions).



Scheme 5.2: The proposed scheme for enantioselective occlusion within racemic β -sheet templates; a majority host species (R_0, S_0) and two minority guest species (R_1, S_1) and (R_2, S_2) of amino acids, all of them provided in racemic proportions. The amino acids of a given chirality attach to sites of the same handedness within the growing β sheet leading to the polymerization of oligomer strands of a uniform chirality (in the alternating row $S-R-S-R-\dots$ fashion). The guest monomers (typically less abundant than the host species) occlude in a random way leading to independent uncorrelated random sequences in each chiral strand. The overall process yields *non-enantiomeric pairs* of homochiral copolymers: mirror symmetry is broken in a stochastic manner.

5.2 Theoretical Methods I: Mass Balance

To address the general setting for the generation of libraries of diastereoisomeric mixtures of peptides as originally proposed by Nery et al. [11], we consider the case where we have a majority amino acid species R_0, S_0 and a given number $m_r, m_s \geq 1$ of minority amino acid species of each chirality (R_1, \dots, R_{m_r}) and (S_1, \dots, S_{m_s}).

¹Note that this does not necessarily imply any net optical activity of solution containing the remaining free chiral monomers.

Since the following calculations are based on chemical equilibrium and detailed balance, if all $(m_r + 1)$ and $(m_s + 1)$ species are supplied in strictly 1:1 racemic proportions (only possible for $m_r \equiv m_s$), we would justifiably expect a racemic outcome, that is, no mirror symmetry breaking. However, we can test the model's ability for chiral amplification by considering unequal initial proportions for the m_r and m_s minority species in solution. That is, does the enantiomeric excess ee increase as a function of chain length, and is it greater than the initial ee of the monomers? i.e., does the model lead to chiral amplification?

From detailed balance, each individual monomer attachment or dissociation reaction is in equilibrium. Then we can compute the equilibrium concentrations of all the (co)-polymers in terms of equilibrium constants K_i for each individual amino acid and the free monomer concentrations. The equilibrium concentration of an S -type copolymer chain of length $n_0 + n_1 + n_2 + \dots + n_{m_s} = N$ made up of n_j molecules type S_j is given by $p_{n_0, n_1, \dots, n_{m_s}}^S = (K_0 s_0)^{n_0} (K_1 s_1)^{n_1} \dots (K_{m_s} s_{m_s})^{n_{m_s}} / K_0$, where $s_j = [S_j]$ [17]. Similarly for the concentration of an R -type copolymer chain of length $n'_0 + n'_1 + n'_2 + \dots + n'_{m_r} = N$ made up of n'_j molecules of type R_j : $p_{n'_0, n'_1, \dots, n'_{m_r}}^R = (K_0 r_0)^{n'_0} (K_1 r_1)^{n'_1} \dots (K_{m_r} r_{m_r})^{n'_{m_r}} / K_0$, where $r_j = [R_j]$. Note that we are considering only copolymers with random sequences such as $R0-R0-R1-R0-R0-R2-R0-\dots$ and $S0-S0-S1-S1-S0-S2-S0-\dots$, but not heterochiral polymers (that is, no sequences involving both the S and R type monomers.) The equilibrium concentration equations we write down $p_{n_0, n_1, \dots, n_{m_r}}^S, p_{n'_0, n'_1, \dots, n'_{m_r}}^R$ implicitly assume the underlying template control².

The number of different S -type copolymers of length l with n_j molecules of type S_j , for $0 \leq j \leq m_s$ species, is given by the multinomial coefficient, hence the total concentration of the S -type copolymers of length l within the β -sheet is given by

$$\begin{aligned} p_l^S &= \sum_{n_0+n_1+\dots+n_{m_s}=l} \binom{l}{n_0, n_1, \dots, n_{m_s}} p_{n_0, n_1, \dots, n_{m_s}}^S \\ &= \frac{1}{K_0} (K_0 s_0 + K_1 s_1 + \dots + K_{m_s} s_{m_s})^l, \end{aligned} \quad (5.1)$$

which follows from the multinomial theorem [18]. From this we can calculate the number of each type S_j of S -monomer present in the S -copolymer of length equal to l , for any $0 \leq j \leq m_s$:

$$\begin{aligned} s_j(p_l^S) &= \sum_{n_0+n_1+\dots+n_{m_s}=l} \binom{l}{n_0, n_1, \dots, n_{m_s}} n_j p_{n_0, n_1, \dots, n_{m_s}}^S = s_j \frac{\partial}{\partial s_j} p_l^S \\ &= \frac{K_j}{K_0} s_j l (K_0 s_0 + K_1 s_1 + \dots + K_{m_s} s_{m_s})^{l-1}. \end{aligned} \quad (5.2)$$

Then we need to know the total amount of the S -type monomers bound within the S -copolymers (in the β -sheet) from the dimer on up to a maximum chain length N . Using Eq.(5.2) for the j th type of amino acid, this is given by

$$s_j(p_{Tot}^S) = \sum_{l=2}^N s_j(p_l^S) \rightarrow \frac{K_j}{K_0} s_j \frac{a(2-a)}{(1-a)^2}, \quad (5.3)$$

²Our above approach assumes that the polymerization reactions are under thermodynamic control. If there are any kinetic effects, they will not be seen as they would contribute to the chain compositions at shorter (finite) time scales. Our aim here is to obtain the compositions at asymptotically long relaxation times, and we thus hypothesize that the dominant pathways are under thermodynamic control.

the final expression holds in the limit $N \rightarrow \infty$ provided that $a = (K_0 s_0 + K_1 s_1 + \dots + K_{m_s} s_{m_s}) < 1$. This must be the case, otherwise the system would contain an infinite number of molecules [17]. Similar considerations hold for the R -sector, and the total amount of R monomers inside R copolymers for the j th amino acid, is given by $r_j(p_{Tot}^R) = \frac{K_j}{K_0} r_j \frac{b(2-b)}{(1-b)^2}$ where $b = (K_0 r_0 + K_1 r_1 + \dots + K_{m_r} r_{m_r}) < 1$. From this we obtain the mass balance equations which hold for both enantiomers S, R of the host and guest amino acids, and is our key result [19]:

$$s_j + \frac{K_j}{K_0} s_j \frac{a(2-a)}{(1-a)^2} = s_{j_{tot}}, \quad r_j + \frac{K_j}{K_0} r_j \frac{b(2-b)}{(1-b)^2} = r_{j_{tot}}. \quad (5.4)$$

These equations express the fact that each type of enantiomer is either free in solution, or is else bound inside a (co)polymer strand within the template.

The problem then consists in the following: given the total concentrations of all the $m_r + 1$ and $m_s + 1$ host plus guest enantiomers $\{s_{j_{tot}}, r_{j_{tot}}\}_{j=0}^{m_s, m_r}$, and the equilibrium constants K_i , we calculate the free monomer concentrations in solution $\{s_j, r_j\}_{j=0}^{m_s, m_r}$ from solving the nonlinear equations Eqs. (5.4). Denote by $s_{0_{tot}} + \dots + s_{m_s_{tot}} + r_{0_{tot}} + \dots + r_{m_r_{tot}} = c_{tot}$ the total system concentration. From the solutions of Eq.(5.4) we can calculate e.g., the equilibrium concentrations of homochiral copolymers $p_{n_0, n_1, \dots, n_{m_s}}^S$ and $p_{n'_0, n'_1, \dots, n'_{m_r}}^R$ of any specific sequence or composition as well as the resultant enantiomeric excess for homochiral chains of length l composed of the host (majority) amino acid:

$$ee_l = \frac{(r_0)^l - (s_0)^l}{(r_0)^l + (s_0)^l}. \quad (5.5)$$

5.3 Theoretical Methods II: Probabilistic approach

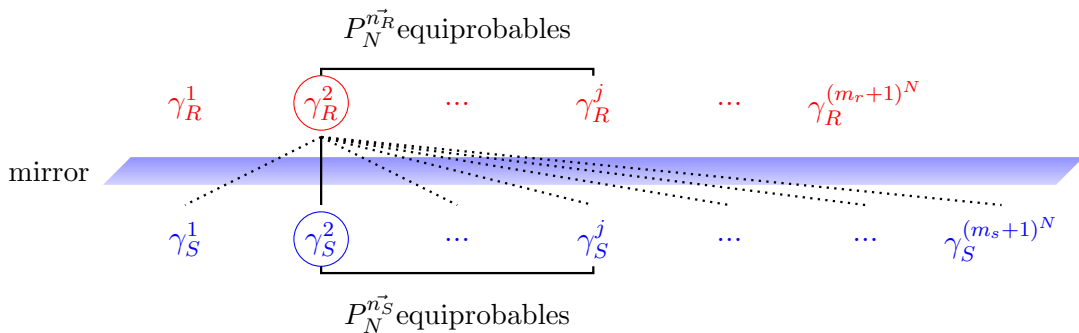
We adopt now a statistical approach for calculating the likelihood for finding *non-enantiomeric pairs* of copolymers formed by the proposed template mechanism. This approach does not require assuming chemical equilibrium.

We first need to specify the length N of the homochiral copolymer chains to be formed, and the number of each minority species or additive m_r, m_s . Thus we consider $(r_0, r_1, r_2, \dots, r_{m_r})$ and $(s_0, s_1, s_2, \dots, s_{m_s})$ whereas $(r, s) \equiv (r_0, s_0)$ denotes both the enantiomers of the majority species. The total number of possible sequences in a chain with N repeat units for each configuration is $(m_r + 1)^N$ and $(m_s + 1)^N$. This most general case is represented in a suggestive pictorial way in Fig. 5.3. This diagram is used to enumerate all possible chiral copolymers that can form in the template, laid out in a linear fashion, the totality of R -copolymers strung out above a "mirror" and the mirror-related S -copolymers directly below it.

Statistical copolymers are those for which the sequence of monomer residues follows a statistical rule. The attachment probability is proportional to the monomer's concentration in solution $[r_j], [s_j]$ times a rate constant that depends on the activation energy E_j for attachment of that specific monomer to the polymer/template thus:

$$p(r_j) \propto A_j \exp(-E_j/kT)[r_j] = w_j[r_j], \quad p(s_j) \propto A_j \exp(-E_j/kT)[s_j] = w_j[s_j]. \quad (5.6)$$

To obtain bona-fide probabilities, these are normalized so that, for $0 \leq j \leq m_r$ and $0 \leq j \leq m_s$ respectively



Scheme 5.3: Homochiral copolymer sequences and their mirror-related sequences. Top row (above the mirror) enumerates all the possible R-type copolymers γ_R of length N that can be made up from m_r different R-type monomers: there are $(m_r + 1)^N$ such chains. Below the mirror: mirror image related S-type homochiral copolymers γ_S made up of S-type monomers. In this example $m_r < m_s$, so there are more S-type copolymers than R-type. Solid vertical line segment links an enantiomeric pair of sequences, the dotted lines represent examples of *non-enantiomeric* pairs of sequences. A given composition typically gives rise to many inequivalent but equiprobable sequences (indicated by the horizontal solid brackets).

$$0 \leq p(r_j) = \frac{w_j[r_j]}{\sum_{k=0}^{m_r} w_k[r_k]} \leq 1, \quad 0 \leq p(s_j) = \frac{w_j[s_j]}{\sum_{k=0}^{m_s} w_k[s_k]} \leq 1, \quad (5.7)$$

which implies $\sum_{k=0}^{m_r} p(r_k) = 1$ and $\sum_{k=0}^{m_s} p(s_k) = 1$.

In writing down Eq.(5.6), there are two implicit assumptions being made: (1) the rate of polymerization is independent of polymer length N (i.e. isodesmic polymerization [20]), and (2), the probability of any given monomer joining a polymer is independent of the existing polymer structure (i.e. first-order Markov process [9])

Define the attachment probability vectors as $\vec{p}_R = \{p(r_0), p(r_1), p(r_2), \dots, p(r_{m_r})\}$ and $\vec{p}_S = \{p(s_0), p(s_1), p(s_2), \dots, p(s_{m_s})\}$, one for the R -monomers, and one for the S -monomers. Note that in the limit when both minority species are absent $m_r, m_s \rightarrow 0$, there will be one unique sequence for each handedness, forming with unit probability.

Chain compositions for the R and S type chains are specified as $n(r_0) + n(r_1) + n(r_2) + \dots + n(r_{m_r}) = N$ and $n(s_0) + n(s_1) + n(s_2) + \dots + n(s_{m_s}) = N$, where $n(r_j)$ and $n(s_j)$ denote the number of times the j -th R and S monomer occur in the corresponding chain, respectively. These are ordered partitions of the integer N . Many different *sequences* can follow from one given composition. By means of the template controlled polymerization mechanism [11], only homochiral chains will be formed, and these can be represented by vectors. For example, a right-handed chain γ_R and its mirror image related *sequence* γ_S are denoted by the vectors

$$\gamma_R = \{r, r, r_1, r, r, r_3, \dots, r\}_N, \quad \gamma_S = \{s, s, s_1, s, s, s_3, \dots, s\}_N. \quad (5.8)$$

The probability to form specific sequences of length N from the compositions is given by the composition probability:

$$p(\gamma_R) = \prod_{j=0}^{m_r} p(r_j)^{n(r_j)}, \quad p(\gamma_S) = \prod_{j=0}^{m_s} p(s_j)^{n(s_j)}. \quad (5.9)$$

In general, there will be many distinct sequences with exactly the same composition-probability Eqs.(5.9), see the horizontal solid line segments in Fig. 5.3.

Probability to form one enantiomeric pair From a given sequence γ_R , we immediately deduce the composition vector $\vec{n}_R = \{n(r_0), n(r_1), \dots, n(r_{m_r})\}$, and the composition vector of the mirror image must be equal to \vec{n}_R , that is $\vec{n}_S = \vec{n}_R \equiv \vec{n}$.

The probabilities to form the specific sequence γ_R and its mirror image γ_S are given by Eq.(5.9) and hence the joint probability to find the *enantiomeric pair* γ_R and γ_S is $P_{pair}(\gamma_R|\gamma_S) = p(\gamma_R)p(\gamma_S)$. This is a function of N , $\min(m_r, m_s)$, \vec{n} , \vec{p}_R and \vec{p}_S .

Probability to form all possible enantiomeric pairs for fixed N For computing the probability of forming all possible enantiomeric pairs, we need to know both m_r and m_s and which one is greater. Without loss of generality we may assume that $m_r \leq m_s$. For fixed N and m_r , the number of distinct compositions of the R type copolymers is given by

$$\#_{N, m_r, \{n_0, n_1, n_2, \dots, n_{m_r}\}} = \binom{m_r + N}{N} = \frac{(m_r + N)!}{N! m_r!}, \quad (5.10)$$

and the number of different sequences that we can form from each individual composition is given by

$$P_N^{\vec{n}_R} \equiv P_N^{n_0 n_1 n_2 \dots n_{m_r}} = \binom{N}{n_0, n_1, n_2, \dots, n_{m_r}} = \frac{N!}{n_0! n_1! n_2! \dots n_{m_r}!}. \quad (5.11)$$

Summing the latter expression over all the possible compositions with fixed N must be equal to the total number of different sequences, that is, we obtain the multinomial theorem [18]: $\sum \binom{N}{n_0, n_1, n_2, \dots, n_{m_r}} = (m_r + 1)^N$.

The net probability we seek to evaluate is $P_{pairs}(N, m_r) = \sum_{\gamma_R} P_{pair}(\gamma_R|\gamma_S)$. This expression is the probability that each and every possible sequence in R and its mirror image sequence in S are formed of fixed length N . For a given composition, all the sequences that can be made therefrom (re-shufflings) are equiprobable. Thus summing over all these possible rearrangements, we arrive at the probability to form chains and their mirror image sequences within one such equiprobable equivalence class, recall $m_r < m_s$:

$$P_N^{\vec{n}_R} p(\gamma_R) p(\gamma_S) = \frac{N!}{n_0! n_1! n_2! \dots n_{m_r}!} \prod_{j=0}^{m_r} p(r_j)^{n(r_j)} \prod_{j=0}^{m_r} p(s_j)^{n(s_j)}. \quad (5.12)$$

Finally, summing this result over all the different compositions, we calculate the net probability to form homochiral chains and their mirror image sequences in the system: i.e., the probability to form all possible enantiomeric pairs. Thus, the probability that mirror symmetry is *not broken* for m additives and N repeat units is given by

$$P_{nobreak}(N, m_r) = P_{pairs}(N, m_r) = \sum_{n_0 + n_1 + n_2 + \dots + n_{m_r} = N} P_N^{\vec{n}} p(\gamma_R) p(\gamma_S). \quad (5.13)$$

Then the probability that mirror symmetry *is* broken for these values of m and N is:

$$\begin{aligned}
P_{break}(N, m_r) &= P_{no\ pairs}(N, m_r) = \sum_{n_0+n_1+n_2+\dots+n_{m_r}=N} P_N^{n_{\vec{R}}} p(\gamma_R)(1 - p(\gamma_S)) \\
&= 1 - \sum_{n_0+n_1+n_2+\dots+n_{m_r}=N} P_N^{n_{\vec{R}}} p(\gamma_R)p(\gamma_S) \\
&= 1 - \left(p_{r_0}p_{s_0} + p_{r_1}p_{s_1} + \dots + p_{r_{m_r}}p_{s_{m_r}} \right)^N = 1 - (\vec{p}_R \cdot \vec{p}_S)^N.
\end{aligned} \tag{5.14}$$

which follows from the multinomial theorem, [18].

From Eqs. (5.7) the monomer attachment probability vectors \vec{p}_R and \vec{p}_S define the faces of two standard or unit m_r and m_s -simplexes [21]. These simplex faces represent the domains of all allowed monomer attachment probabilities, see Ref. [22] for a detailed explanation and graphic representations. Typically, copolymers will be formed with lengths ranging from the dimer, trimer, etc. on up to a maximum number of repeat units N [11, 12, 15]. The above arguments apply to any value of N , thus the probability $P_{break}^{\leq N}(m)$ to break mirror symmetry in a system containing a spectrum of chain lengths $2 \leq n \leq N$ is given as follows,

$$P_{break}^{\leq N}(m_r) = \frac{1}{N-1} \sum_{n=2}^N P_{break}(n, m_r) = 1 - \frac{(\vec{p}_R \cdot \vec{p}_S)^2 (1 - (\vec{p}_R \cdot \vec{p}_S)^{N-1})}{(N-1)(1 - \vec{p}_R \cdot \vec{p}_S)}, \tag{5.15}$$

and satisfies $\lim_{\vec{p}_R \cdot \vec{p}_S \rightarrow 1} P_{break}^{\leq N}(m) = 0$ when the two occlusion probability vectors are parallel.

5.4 Results

Additives of only one handedness

Mass balance Following the model introduced in Section 5.2, the three monomer case originally treated [11] corresponds to $m_r = 0$ and $m_s = 1$ (the system contains R_0, S_0 and only the enantiomer S_1 of the guest species). We consider stepwise additions and dissociations of single monomers from one end of the (co)polymer chain, considered as a strand within the β -sheet, see Fig. 5.2. It is reasonable to regard the β -sheet in equilibrium with the free monomer pool³ [23]. For a single guest, we drop numbered indices and denote the majority host species and concentrations by $r = [R], s = [S]$ and the minority guest with a prime: $s' = [S']$. In Fig. 5.1a we plot the numerical results obtained from calculating the enantiomeric excess ee as a function of chain length l , from Eq.5.5, for the three starting compositions of the monomer crystals as reported [11]. The only quantities required for this are the solutions of r and s obtained from solving the set of equations Eq.(5.4). For strictly illustrative purposes only, we set the equilibrium constants to be the same for

³Ref. [23] reports a stochastic simulation of two concurrent orthogonal processes: 1) an irreversible condensation of activated amino acids and 2) reversible formation of racemic β -sheets of alternating homochiral strands. The two steps taken together comprise a two-dimensional formulation of the problem. These architectures lead to the formation of chiral peptides whose isotacticity increases with length (we assume as given the prior formation of the initial templates or β -sheets, and are concerned exclusively with the subsequent enantioselective random polymerization reactions (step (2)))

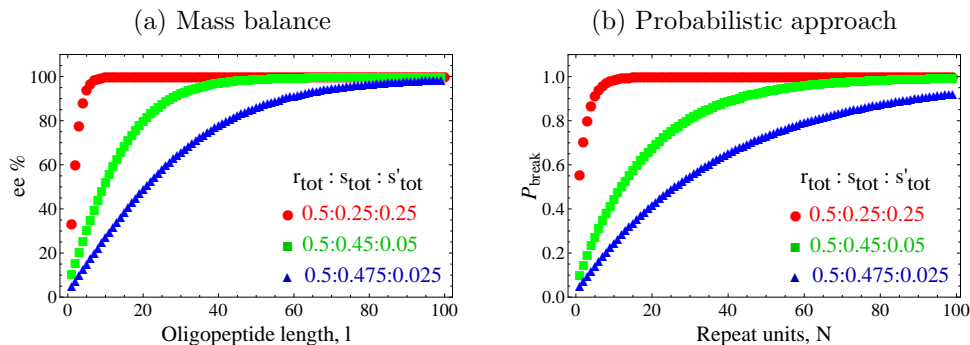


Figure 5.1: (a) Calculated ee values from solving Eqs. (5.4) for $m = 1$ guest monomer and three different starting monomer compositions (in relative proportions) $r_{tot} : s_{tot} : s'_{tot} = 0.5 : 0.25 : 0.25$ (filled circles), $0.5 : 0.45 : 0.05$ (squares) and $0.5 : 0.475 : 0.025$ (triangles) for the equilibrium constant $K_0 = K_1 = 1000M^{-1}$ and the total monomer concentration $c_{tot} = 10^{-3}M$. (b) Probability to break mirror symmetry, $P_{break}^{\leq N}(m)$, for the three different starting composition $f_r : f_s : f_{s'}$ of the three component case ($m_r = 0$ and $m_s = m = 1$) as a function of repeat units N . Compare to Fig. 13 of Nery et al. [11].

both host and guest monomers $K_1 = K_0 \equiv K = 1000M^{-1}$, the total initial concentration, $c_{tot} = 10^{-3}M$; the initial fractions of each component are denoted by $f = \{f_r, f_s, f_{s'}\}$ and obey $f_r + f_s + f_{s'} = 1$. The starting composition of the mixture is $c_{tot} = r_{tot} + s_{tot} + s'_{tot}$, and the total amount of each component is: $r_{tot} = c_{tot} * f_r$, $s_{tot} = c_{tot} * f_s$, and $s'_{tot} = c_{tot} * f_{s'}$. We can appreciate the induced symmetry breaking mechanism [11] from the behavior of ee_l .

For the first case $f_r : f_s : f_{s'} = 0.5 : 0.25 : 0.25$, mirror symmetry is broken for almost all the chain lengths, even for small values of l ; this is due to the equal starting fractions of the majority s_{tot} and the guest s'_{tot} monomer species of the same chirality: the large amount of guest is the reason for these large values of ee . For the second case $f_r : f_s : f_{s'} = 0.5 : 0.45 : 0.05$, the starting fraction of the majority species, s'_{tot} , is almost 10 times ($0.45/0.05 = 9$) greater than that of the guest, s' , so for the enantiomeric excess to be greater than 60% the chain length must be at least $l = 13$, and for obtaining an ee of 80%, the chain length must be at least $l = 20$. Finally, for the third case, $f_r : f_s : f_{s'} = 0.5 : 0.475 : 0.025$, the starting fraction of the majority species, s_{tot} , is almost 20 times ($0.475/0.025 = 19$) greater than that of the guest, s'_{tot} , thus the enantiomeric excess for each chain length is expected to be much less than for the two previous cases. For the three cases, an increase of the ee_l is observed (for all l) when increasing the starting fraction of the guest species, s'_{tot} . When s'_{tot} is comparable to s_{tot} , while maintaining the overall proportion R -type: S -type=1:1, then symmetry breaking is ensured to be $> 40\%$ for all $l > 5$.

Probabilistic approach Consider first the general case $m_r = 0$ and we set $m_s = m$. Clearly, there is only one possible composition (and hence, sequence) that can be formed in R : namely $\gamma_R = \{r, r, \dots, r\}_N$, and this forms with unit probability: $p(\gamma_R) = 1$. Thus, the probability that mirror symmetry is *not broken* for m types of S -additives and for N repeat units is given by Eq.(5.13), which simplifies to give:

$$P_{nobreak}(N, m) = P_{pairs}(N, m) = p(\gamma_R)p(\gamma_S) = p(s_0)^N. \quad (5.16)$$

Then the probability that mirror symmetry *is* broken for these values of m and N is:

$$P_{break}(N, m) = P_{nopairs}(N, m) = p(\gamma_R)(1 - p(\gamma_S)) = 1 - p(s_0)^N. \quad (5.17)$$

If the number of S -type additives goes to zero, $m_s \rightarrow 0$, then $p(s_0) \rightarrow 1$ and then mirror symmetry is maintained with absolute certainty.

The probability of breaking mirror symmetry depends only on $p(s_0)$: the attachment probability of the S -enantiomer of the majority species. The minimal probability of breaking symmetry $P_{break}(N, m) = 0$ is obtained for $p(s_0) = 1$, in this case, there are no guests, only the majority species S_0 , so we recover the case in which additives are supplied in racemic proportions, and moreover, no guests are added. The maximum probability of breaking mirror symmetry $P_{break}(N, m) = 1$ is obtained for $p(s_0) = 0$. In this case, the majority species in S , is absent, thus no possible enantiomeric pairs can be formed. The probability $P_{break}^{\leq N}(m)$ to break mirror symmetry in a system containing a spectrum of chain lengths $2 \leq n \leq N$ Eq.(5.15) reduces to

$$P_{break}^{\leq N}(m) = 1 - \frac{p(s_0)^2(1 - p(s_0)^{N-1})}{(N-1)(1 - p(s_0))}, \quad (5.18)$$

and satisfies $\lim_{p(s_0) \rightarrow 1} P_{break}^{\leq N}(m) = 0$ when no majority species of the S -type is supplied.

Following Eq.(5.18) we can calculate $P_{break}^{\leq N}(m)$ for the three different starting compositions considered before, exemplary numerical results are shown in Fig. 5.1b showing the effect of varying the relative concentrations of all the monomers and the activation energy (we vary w'_s) of the guest monomer s' (see Ref. [22] for numerical values). The curves for P_{break} are qualitatively similar to those of the percent ee in Fig. 5.1a.

Racemic additives

Mass balance By way of one further example, we carry out a similar analysis for the case of four monomers, this time for two majority R, S and two minority amino acids: R', S' ($m_r = m_s = m$). Fig. 5.1a and Fig. 5.2 clearly demonstrate that the higher (lower) is the initial degree of chiral asymmetry, characterized by r_{tot}/s_{tot} in the former and r'_{tot}/s'_{tot} in the latter, the higher (lower) is the final asymmetry. Thus, rather than symmetry breaking *per se*, we are observing the model's capacity for asymmetric amplification, as stated at the beginning of Section 5.2. Nevertheless, effects closer to a symmetry breaking effect can be appreciated by looking at the average chain lengths for unequal equilibrium constants (see Ref. [19] or Ref. [22]).

Probabilistic approach From Eq.(5.5) we calculate the ee_l for the different chain lengths l for three different starting compositions. In Fig. 5.2 we show the numerical results obtained from the solutions of the set of equations Eq.(5.4) and Eq.(5.5), for $K_0 = K_1 = 1000M^{-1}$ and $s_{tot} + s'_{tot} + r_{tot} + r'_{tot} = 10^{-3}M$. When the enantiomeric species are provided in ideally racemic proportions, the probability that mirror symmetry is broken for given values of the chain length N and number of species m can be expressed succinctly as

$$P_{break}(N, m) = 1 - \|\vec{p}\|^{2N}, \quad (5.19)$$

that is, one minus the squared-modulus of the probability attachment vector \vec{p} , raised to the chain length. Thus, for fixed N , in order to maximize the probability that mirror symmetry be broken, we should prepare the chemical system so that all m additives and the majority species have equally shared mole fractions. For any other values, but excluding $m + 1$, then $\|\vec{p}\| < 1$, hence the probability to break mirror symmetry increases with chain length N and/or with increasing number of additives m , provided these are supplied with small mole fractions.

Finally, if $\|\vec{p}\| = 1$, so $P_{break}(N, m) = 0$, and mirror symmetry is maintained with absolute certainty for all N .

If we increase the mole fraction of any one of the additives in excess, the tables are turned, and the majority and minority species interchange their roles: excessive amounts of any additive tend to reduce the probability for breaking mirror symmetry. Eq.(5.15) now simplifies to give

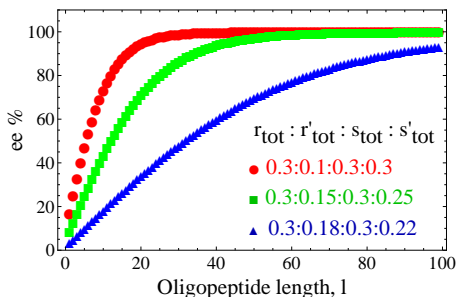


Figure 5.2: Calculated ee values for three different starting monomer compositions $r_{tot} : r'_{tot} : s_{tot} : s'_{tot}$ for $K_0 = K_1 = 1000M^{-1}$ and $c_{tot} = 10^{-3}M$.

and satisfies $\lim_{\|\vec{p}\| \rightarrow 1} P_{break}^{\leq N}(m) = 0$. As expected, we find increasing probability for symmetry breaking as N and/or m increase. And the probability to break mirror symmetry is strictly zero when there are no additives: $P_{break}^{\leq N}(m = 0) = 0$. The cases with two majority species r and s and two guests, r' and s' , with starting fractions $f_r : f_{r'} : f_s : f_{s'}$, as con-

sidered in the first section of the present chapter, is a case of racemic additives where $m_r = m_s = 1$. Following Eq.(5.20) we can calculate $P_{break}^{\leq N}(m)$ for the three different starting compositions considered before. The results (not shown) are qualitatively very similar to those shown in the previous tables.

5.5 Conclusions

- The proposed experimental mechanism [11] leads to the formation of homochiral copolymers with random sequences of the majority and minority amino acids. Given the implications of the experimental mechanism, we have provided two independent and complementary theoretical approaches to the problem. Both these approaches provide further quantitative insights into the template-controlled induced desymmetrization mechanisms advocated by Lahav and coworkers [10–16].

- Our two theoretical models invoke the underlying template control in that they do not allow for any heterochiral oligomers to form. The sequence of the host and guest amino acids within the homochiral peptides assembles in a completely random fashion, in accord with the experiments [11]. This sequence randomness is captured by both the model based on chemical equilibrium and by the second model based on the monomer occlusion probabilities.

• The statistical/combinatorial effects do lead to a stochastic mirror symmetry breaking effect. The symmetry breaking in these experiments arises from combinatorics, not from spontaneous (bifurcation) phenomena. These stochastic/statistical/combinatorial effects are not due to the inherent tiny chiral fluctuations present in all real chemical systems [24–26] but are due rather to the random occlusion of host and guest amino acids by the chiral sites of the template: the mechanisms proposed here work even for ideally racemic mixtures. Mirror symmetry is broken in the sequences, as non-enantiomeric pairs of oligomers are formed. The solution of free monomers can nevertheless be optically inactive. The symmetry breaking is to be found in the template, or β -sheet, but not (necessarily) in the solution.

Bibliography

- [1] Bada, J. and Miller, S. *Biosystems* **20**, 21 (1987).
- [2] Avetisov, V., Goldanskii, V. and Kuzmin, V. *Dokl. Akad. Nauk USSR* **115**, 282 (1985).
- [3] Goldanskii, V., Avetisov, V. and Kuzmin, V. *FEBS Lett.* **207**, 181 (1986).
- [4] Avetisov, V. and Goldanskii, V. *Proc. Natl. Acad. Sci. USA.* **93**, 11435 (1996).
- [5] Orgel, L. *Nature* **358**, 203 (1992).
- [6] Blocher, M., Hitz, T. and Luisi, P. *Helv. Chim. Acta* **84**, 842 (2001).
- [7] Hitz, T., Blocher, M., Walde, P. and Luisi, P. L. *Macromolecules* **34**, 2443 (2001).
- [8] Guijarro, A. and Yus, M. *The Origin of Chirality in the Molecules of Life* (RSC Publishing, Cambridge, 2009), 1st ed.
- [9] Weissbuch, I., Illos, R. A., Bolbach, G. and Lahav, M. *Accounts Chem. Res.* **42**, 1128 (2009).
- [10] Zepik, H., Shavit, E., Tang, M., Jensen, T., Kjaer, K., Bolbach, G., Leiserowitz, L., Weissbuch, I. and Lahav, M. *Science* **295**, 1266 (2002).
- [11] Nery, J., Bolbach, G., Weissbuch, I. and Lahav, M. *Chem. Eur. J.* **11**, 3039 (2005).
- [12] Nery, J., Eliash, R., Bolbach, G., Weissbuch, I. and Lahav, M. *Chirality* **19**, 612 (2007).
- [13] Rubinstein, I., Eliash, R., Bolbach, G., Weissbuch, I. and Lahav, M. *Angew. Chem. Int. Edit.* **46**, 3710 (2007).
- [14] Rubinstein, I., Clodic, G., Bolbach, G., Weissbuch, I. and Lahav, M. *Chem-Eur. J.* **14**, 10999 (2008).
- [15] Illos, R. A., Bisogno, F. R., Clodic, G., Bolbach, G., Weissbuch, I. and Lahav, M. *J. Am. Chem. Soc.* **130**, 8651 (2008).
- [16] Illos, R., Clodic, G., Bolbach, G., Weissbuch, I. and Lahav, M. *Origins Life Evol. B.* **40**, 51 (2010).
- [17] Markvoort, A., ten Eikelder, H., Hilbers, P., de Greef, T. and Meijer, E. *Nat. Commun.* (2011).
- [18] Digital library of mathematical functions. Tech. Rep., National Institute of Standards and Technology, <http://dlmf.nist.gov/26>.
- [19] **Blanco, Celia** and Hochberg, D. Induced mirror symmetry breaking via template-controlled copolymerization: theoretical insights. *Chem. Commun.* **48**, 3659–3661 (2012).
- [20] De Greef, T. F. A., Smulders, M. M. J., Wolffs, M., Schenning, A. P. H. J., Sijbesma, R. P. and Meijer, E. W. *Chem. Rev.* **109**, 5687 (2009).
- [21] W. Rudin. *Principles of Mathematical Analysis* (McGraw-Hill, New York, 1976).
- [22] **Blanco, Celia** and Hochberg, D. Models for mirror symmetry breaking via β -sheet-controlled copolymerization:(i) mass balance and (ii) probabilistic treatment. *J. Phys. Chem. B* **116**, 13953–13967 (2012).
- [23] Wagner, N., Rubinov, B. and Ashkenasy, G. *ChemPhysChem* **12**, 2771 (2011).
- [24] Mills, W. *J. Soc. Chem. Ind.* **10**, 75 (1932).
- [25] Mislow, K. *Collect. of Czech. Chem. C.* **68**, 849 (2003).
- [26] Siegel, J. *Chirality* **10**, 24 (1998).

Conclusions

- In the Frank model, strong mutual inhibition is required to amplify the initial ee to large values, and temporary chiral excursions are observed for closed and semi-open systems and can be explained through phase space analysis, stability analysis and numerical simulations. Such chiral excursions may be experimentally observed and could be mistaken for a transition to a chiral state. In the system closed to matter flow and subject to constraints imposed by micro-reversibility, total mass is conserved and temporary symmetry breaking can occur but never permanent symmetry breaking: SMSB is a kinetically controlled emergence of chirality. The system eventually reaches chemical equilibrium, where the racemic state is the only stable one.
- The LES model may lead to SMSB (as a stationary final state) whether in a cyclic network if the reverse reaction of the nonenantioselective autocatalysis is driven by an external reagent, or in a two-compartment model if both autocatalytic reactions are spatially separated at different temperatures in different compartments but coupled under the action of a continuous internal flow. In both systems, total mass is conserved, but permanent symmetry breaking can occur (the external reagent and the temperature gradient respectively ensure conditions that keep the system out of chemical equilibrium).
- A strong but temporary chiral amplification can take place in both reversible models of chiral polymerization and copolymerization, closed to matter flow and subject to constraints imposed by micro-reversibility. Total mass is conserved in both systems, which necessary will reach chemical equilibrium, where the racemic state is the only stable one. The chain length dependent enantiomeric excesses tend to exhibit a damped oscillatory behavior before the onset of final racemization and the rate of entropy production per unit volume exhibits a peak value either before or near the vicinity of the chiral symmetry breaking transition.
- The copolymerization model provides robustness of the fits with respect to differing initial enantiomeric excesses and qualitatively reproduces its temporal behavior. The correlation between calculated and experimental relative abundances is greater for the initially non-racemic situations: our simple copolymerization scheme does provide an accurate course-grained description of the lattice-controlled polymerization reported by Lahav and coworkers.
- In all the five cases: the Frank model, the LES model (considering the necessary conditions above mentioned), the polymerization model and the copolymerization model, the system can evolve, for certain parameters, towards a chiral stationary state, amplifying a tiny initial enantiomeric excess (even lower than the calculated ee_{st}). It is worth mentioning that, for the polymerization and copolymerization models, this tiny initial enantiomeric excess can not only be amplified but transmitted to the longest chains (note that initial conditions assume an initial imbalance in the monomers while the rest of initial concentrations are set to zero).
- The two theoretical approaches for the experimental formation of homochiral copolymers provide further quantitative insights into the template-controlled induced desymmetrization mechanisms advocated by Lahav and coworkers. Both models invoke the underlying template control. The symmetry breaking in these experiments arises from combinatorics, not from spontaneous (bifurcation) phenomena. These stochastic/statistical/combinatorial effects are not due to the inherent tiny chiral fluctuations present in all real chemical systems but are due rather to the random occlusion of host and guest amino acids by the chiral sites of the template: the mechanisms proposed here work even for ideally racemic mixtures.

Conclusiones

- En el modelo de Frank, es necesaria una fuerte inhibición mutua para amplificar el ee inicial hasta valores altos. Se observan excursiones quirales en sistemas cerrados y semiabiertos que pueden ser explicadas mediante análisis del espacio de fases, análisis de estabilidad y simulaciones numéricas. Tales excursiones quirales pueden ser observadas experimentalmente y ser confundidas con una transición al estado quiral. En un sistema cerrado al flujo de materia y sujeto a ligaduras impuestas por microreversibilidad, la masa total se conserva y puede tener lugar ruptura temporal de simetría, pero nunca de manera permanente: SMSB es emergencia de quiralidad controlada cinéticamente. El sistema finalmente alcanza el equilibrio químico, siendo el estado racémico el único estado estable posible.
- El modelo de LES puede dar lugar a SMSB tanto en un mecanismo cíclico si la reacción inversa de la autocatálisis no enantioselectiva es conducida por un agente externo, o en un modelo de dos compartimentos si ambas reacciones autocatalíticas están separadas a diferentes temperaturas en diferentes compartimentos pero unidos bajo la acción de un flujo interno constante. En ambos sistemas, la masa se conserva, pero la ruptura permanente de simetría puede tener lugar (el agente externo y el gradiente de temperatura aseguran condiciones que mantienen al sistema fuera del equilibrio químico).
- Una amplificación quiral fuerte, pero temporal, puede tener lugar en los modelos reversibles de polimerización y copolimerización quiral cerrados al flujo de materia y sujetos a ligaduras impuestas por microreversibilidad. La masa total se conserva en ambos sistemas, que necesariamente alcanzan equilibrio químico, en el cuál, el estado racémico es el único estable. El exceso enantiomérico en función de la longitud de las cadenas muestra un comportamiento oscilatorio amortiguado antes de que comience la racemización final, y la tasa de producción de entropía por unidad de volumen muestra un pico o bien antes, o bien cerca de la transición de ruptura de simetría quiral.
- El modelo de copolimerización ofrece ajustes consistentes con respecto a diferentes excesos iniciales, y reproduce cualitativamente su comportamiento temporal. La correlación entre las abundancias relativas calculadas y las experimentales aumenta para situaciones iniciales no racémicas: nuestro esquema sencillo de copolimerización quiral ofrece una descripción apta de la polimerización por control de plantillas descrito por Lahav y su grupo.
- En los cinco casos: el modelo de Frank, el modelo de LES (considerando las condiciones necesarias mencionadas más arriba), el modelo de polimerización y el de copolimerización quiral, el sistema puede evolucionar, para determinados parámetros, hacia un estado estacionario quiral, amplificando el pequeño exceso enantiomérico inicial (incluso por debajo del calculado, ee_{st}). Cabe mencionar que, para los modelos de polimerización y copolimerización quiral, ese pequeño exceso enantiomérico inicial, no sólo es amplificado, sino también transmitido a las cadenas más largas (las condiciones iniciales asumen un desequilibrio inicial en los monómeros, mientras que el resto de concentraciones iniciales son nulas)
- Los dos enfoques teóricos propuestos para la formación experimental de copolímeros homquirales ofrecen un entendimiento cuantitativo de los mecanismos de desimetrización inducida propuestos por Lahav y su grupo. Ambos modelos asumen el control de plantilla subyacente. La ruptura de simetría en estos experimentos surge de la combinatoria, no de fenómenos espontáneos (bifurcaciones). Estos efectos estocásticos/estadísticos/combinatorios no son debido a pequeñas fluctuaciones quirales presentes en todos los sistemas químicos, sino más bien a la oclusión aleatoria de amino ácidos huéspedes y anfitriones en los sitios quirales de la plantilla: los mecanismos propuestos aquí funcionan incluso para mezclas perfectamente racémicas.



Temporary mirror symmetry breaking and chiral excursions in open and closed systems

Celia Blanco, Michael Stich, David Hochberg*

Centro de Astrobiología (CSIC-INTA), Ctra de Ajalvir km 4, 28850 Torrejón de Ardoz, Madrid, Spain

ARTICLE INFO

Article history:

Received 30 December 2010

In final form 13 February 2011

Available online 16 February 2011

ABSTRACT

The reversible Frank model is capable of amplifying the initial small statistical deviations from the idealized racemic composition. This temporary amplification can be interpreted as a chiral excursion in a dynamic phase space. It is well known that if the system is open to matter and energy exchange, a permanently chiral state can be reached asymptotically, while the final state is necessarily racemic if the system is closed. In this work, we combine phase space analysis, stability analysis and numerical simulations to study the initial chiral excursions and determine how they depend on whether the system is open, semi-open or closed.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The Frank model [1] has been extensively invoked to justify theoretically the emergence of biological homochirality [2,3], and is usually analyzed as a reaction network in open systems (matter and energy are exchanged with the surroundings) composed of an irreversible enantioselective autocatalysis coupled to an irreversible mutual inhibition reaction between the product enantiomers. The model shows how homochirality is achieved as a stationary state when the mutual inhibition product (the heterodimer) is removed from the system and when the concentration of the achiral substrate is held constant. By contrast, for reversible transformations and when the mutual inhibition product remains in the system, the final stable state can only be the racemic one. As a consequence, a thermodynamically controlled spontaneous mirror symmetry breaking (SMSB) cannot be expected to take place. In particular, SMSB is not expected for reversible reactions taking place in systems closed to matter and energy flow.

Nevertheless, as was recently demonstrated [4] for systems closed to matter flow, the Frank model is a prime candidate for the fundamental reaction network necessary for reproducing the key experimental features reported on absolute asymmetric synthesis in the absence of any chiral polarization [5]. Most importantly, when reversible steps in all the reactions are allowed it is capable of [4] (i) amplification of the initially tiny statistical enantiomeric excesses from $ee \sim 10^{-8}\%$ to practically 100%, leading to (ii) long duration chiral excursions or chiral pulses away from the racemic state at nearly 100% ee , followed by, (iii) the final approach to the stable racemic state for which $ee = 0$, i.e., mirror symmetry

is recovered permanently. To understand this temporary asymmetric amplification is important because the racemization time scale can be much longer than that for the complete conversion of the achiral substrate into enantiomers.

Long duration chiral excursions have also been reported recently in closed chiral polymerization models with reversible reactions [6] where constraints implied by microreversibility have been taken into account. These results are important because they suggest that temporary spontaneous mirror symmetry breaking in experimental chiral polymerization can take place, and with observable and large chiral excesses without the need to introduce chiral initiators [7] or large initial chiral excesses [8].

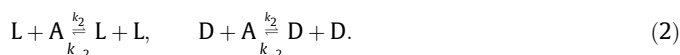
The purpose of this Letter is to elucidate the nature of these chiral excursions by combining the information provided by phase plane portraits, numerical simulation and linear stability analysis. We consider the Frank model, this being the most amenable to such types of analysis and because it is the ‘common denominator’ of numerous more elaborate theoretical models of SMSB [3].

The reaction scheme consists of a straight non-catalyzed reaction Eq. (1), an enantioselective autocatalysis Eq. (2), where A is a prechiral starting product, and L and D are the two enantiomers of the chiral product. We also assume reversible heterodimerization step in Eq. (3), where LD is the achiral heterodimer. The k_i denote the reaction rate constants. In the following, we give the reaction steps in detail.

Production of chiral compound:



Autocatalytic amplification:



* Corresponding author.

E-mail address: hochberg@cab.inta-csic.es (D. Hochberg).

Hetero-dimerization:



We assume the feasibility of the reverse reaction for all the steps. Focusing our attention on chiral excursions, we make a careful distinction between open, semi-open or fully closed systems. These system constraints are crucial for determining both the intermediate and the asymptotic final states of the chemical system.

2. Open system

2.1. Rate equations

We first consider the original Frank scenario [1]. There, steady and stable chiral states can be achieved, since the system is permanently held out of equilibrium. See [9] for more details. An important question is, can the system support chiral excursions? That is, pass through temporary chiral states before ending up in the final racemic state?

In the original Frank model there is an incoming flow of achiral compound A and elimination of the heterodimer LD from the system. A convenient way to account for the inflow of achiral matter is to assume that the concentration of the prechiral component [A] is constant, and then we need not write the corresponding kinetic equation for it. For the outflow the heterodimer leaves the system at a rate $\bar{\gamma}$. We assume that the heterodimer formation step is irreversible, and set $k_{-5} = 0$. Note, the elimination of LD from the system can actually be neglected as long as the hetero-dimerization step is irreversible [3]. We retain this outflow however since it is needed to obtain stationary asymptotic values of all three concentrations [L], [D] and [LD]; see the fixed points below. So with [A] = const and replacing Eq. (3) by



we obtain the rate equations

$$\frac{d}{dt}[L] = k_1[A] + (k_2[A] - k_{-1})[L] - k_{-2}[L]^2 - k_5[L][D], \quad (6)$$

$$\frac{d}{dt}[D] = k_1[A] + (k_2[A] - k_{-1})[D] - k_{-2}[D]^2 - k_5[D][L], \quad (7)$$

$$\frac{d}{dt}[LD] = k_5[L][D] - \bar{\gamma}[LD]. \quad (8)$$

The key variable throughout is the chiral polarization

$$\eta = \frac{[L] - [D]}{[L] + [D]}, \quad (9)$$

also called enantiomeric excess *ee*, which obeys $-1 \leq \eta \leq 1$ and which represents the order parameter for mirror symmetry breaking.

In order to simplify the analysis, we define a dimensionless time parameter $\tau = (k_2[A] - k_{-1})t$ and dimensionless concentrations that scale as $[\tilde{L}] = k_5(k_2[A] - k_{-1})^{-1}[L]$, $[\tilde{D}] = k_5(k_2[A] - k_{-1})^{-1}[D]$, $[\tilde{LD}] = k_5(k_2[A] - k_{-1})^{-1}[LD]$. It is convenient to define the sums and differences of concentrations: $\chi = [\tilde{L}] + [\tilde{D}]$, $y = [\tilde{L}] - [\tilde{D}]$, and for the heterodimer put $P = [\tilde{LD}]$. The chiral polarization $\eta = y/\chi$ remains unchanged.

In terms of the new variables, Eqs. (6)–(8) read

$$\frac{d\chi}{d\tau} = 2u + \chi - \frac{1}{2}(g+1)\chi^2 - \frac{1}{2}(g-1)\chi^2\eta^2, \quad (10)$$

$$\frac{d\eta}{d\tau} = \eta(1 - g\chi) - \frac{\eta}{\chi} \left(\frac{d\chi}{d\tau} \right), \quad (11)$$

$$\frac{dP}{d\tau} = \frac{1}{4}\chi^2(1 - \eta^2) - \gamma P. \quad (12)$$

The dimensionless parameters appearing here are:

$$u = \frac{k_1 k_5 [A]}{(k_2 [A] - k_{-1})^2}, \quad g = \frac{k_{-2}}{k_5}, \quad \gamma = \frac{\bar{\gamma}}{(k_2 [A] - k_{-1})}. \quad (13)$$

The system is described by three equations (10)–(12). Since P does not enter into the equations for χ and η , the equations decouple and the dynamical system to study is effectively two-dimensional and so the appearance of SMSB cannot depend on whether the heterodimer is removed from the system when $k_{-5} = 0$, although the fixed points will (see Section 2.3).

2.2. Phase plane and linear stability analysis

In the phase space of the dynamical system defined by Eqs. (10) and (11) there are curves with a special significance. These are the nullclines defined by

$$\frac{d\chi}{d\tau} = 0, \quad (14)$$

$$\frac{d\eta}{d\tau} = 0. \quad (15)$$

The intersections of these curves give the possible steady states (or fixed points) of the system. The condition $d\chi/d\tau = 0$ leads to two curves

$$\chi_{\pm}^{(1)} = \frac{1 \pm \sqrt{1 + 4u(g(1 + \eta^2) + 1 - \eta^2)}}{g(1 + \eta^2) + 1 - \eta^2}, \quad (16)$$

while $d\eta/d\tau = 0$ implies the three curves

$$\eta = 0, \quad \chi_{\pm}^{(2)} = \pm \sqrt{\frac{4u}{1 - g + (g-1)\eta^2}}. \quad (17)$$

For $u > 0$ the solutions denoted χ_- correspond to negative total enantiomer concentrations so we discard them. The physically acceptable nullclines are plotted in Figure 1. Which of the two different intersection configurations is obtained depends only on the single parameter g . We emphasize that despite the similar appearance, the nullcline graphs should not be confused with the classic bifurcation diagrams that have been discussed often in the past [10–12].

2.3. Fixed points and stability

The system has several steady states: besides an unphysical state that we disregard, there is a Z_2 pair of chiral solutions Q_{\pm} , and a racemic state R :

$$R = \left(P = \frac{2(g+1)u + \sqrt{4(g+1)u + 1} + 1}{2(g+1)^2\gamma}, \right. \\ \left. \chi = \frac{1 + \sqrt{4(g+1)u + 1}}{g+1}, y = 0 \right), \quad (18)$$

$$Q_{\pm} = \left(P = \frac{u}{\gamma - g\gamma}, \chi = \frac{1}{g}, y = \pm \frac{\sqrt{((g-1)/g^2) + 4u}}{\sqrt{g-1}} \right). \quad (19)$$

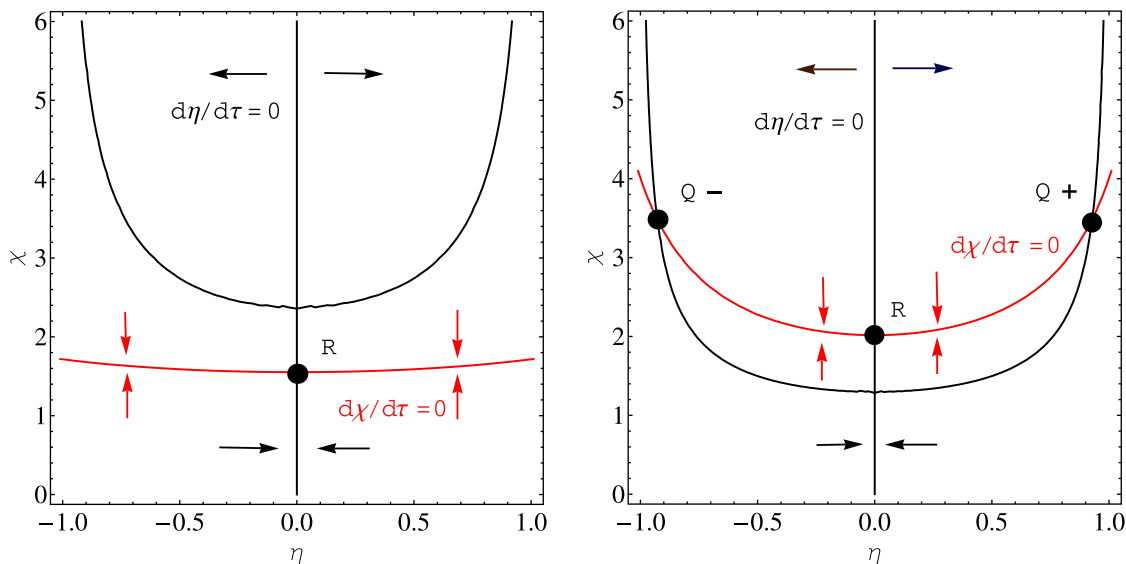


Figure 1. Nullclines for the open system (10) and (11). These curves correspond to $u = 0.3$ implying $g_{crit} = 0.59$. The $d\eta/d\tau = 0$ and $d\chi/d\tau = 0$ nullclines are plotted in black and red, respectively. The black (red) arrows indicate the regions of phase space where $\eta(\chi)$ increases or decreases. Left: $g = 0.79 > g_{crit}$. The nullclines intersect in only the one point R representing the asymptotic stable racemic state. Right: $g = 0.29 < g_{crit}$. In this case there are three intersections, Q_{\pm} and R representing the Z_2 equivalent stable chiral states and the unstable racemic state, respectively. From Eq. (22) the enantiomeric excess at Q_{\pm} is $\eta = \pm 0.93$.

The associated eigenvalues are given by [9]:

$$\lambda_{1,2,3}(R) = \left(-\sqrt{4(g+1)u+1}, \frac{1-g\sqrt{4(g+1)u+1}}{g+1}, -\gamma \right), \quad (20)$$

$$\lambda_{1,2,3}(Q_{\pm}) = \left(\frac{-\sqrt{16g^3u+4g^2-4g+1}-1}{2g}, \frac{\sqrt{16g^3u+4g^2-4g+1}-1}{2g}, -\gamma \right). \quad (21)$$

Note that $\lambda_1(R) < 0$ and $\lambda_1(Q) < 0$ are always negative whereas $\lambda_2(R) > 0$ and $\lambda_2(Q) < 0$ for $g < g_{crit}$, otherwise $\lambda_2(R) < 0$ and $\lambda_2(Q) > 0$ for $g > g_{crit}$, where $g_{crit} = (\sqrt{1+16u}-1)/8u$ is the critical value for this parameter. Note that $g_{crit}(u) \leq 1$ for all $u \geq 0$. For small u we can write $g_{crit} = 1-4u$; while for large u , $g_{crit} \rightarrow 1/2u^{1/2}$. Thus the direct chiral monomer production step ($\propto k_1$ in (1)) tends to racemize the system leading to final η values strictly less than unity:

$$\eta = \pm \sqrt{1 - \frac{4ug^2}{1-g}}, \quad (22)$$

which holds when $g < g_{crit}$. The chiral monomer production step thus reduces the range of g for which stable mirror symmetry breaking can occur, and the chiral solutions are no longer 100% chiral.

Figure 2 shows the temporal evolution of the L and D chiral monomers starting from an extremely dilute total enantiomer concentration and the very small statistical chiral deviations from the ideal racemic composition. The right hand side of this figure shows the evolution in terms of the quantities χ and η . Note the mirror symmetry breaking signalled by η . Are there chiral excursions found in the open system model? A chiral excursion holds when the enantiomeric excess η departs from a small initial value, evolves to some maximum absolute value and then decays to the final value of zero. To ensure a final racemic state we must set $g > g_{crit}$ but then we find no numerical evidence for such temporary chiral excursions. This can be understood qualitatively from inspection of left hand side of Figure 1. The initial conditions (dilute chiral monomer concentration and statistical chiral fluctuation) corresponds to a initial point located at tiny values of χ and close to the vertical nullcline, well below the point labeled as R . The system is attracted to the black curve and moves up the curve to R . In this situation, it is impossible for the chiral excess to

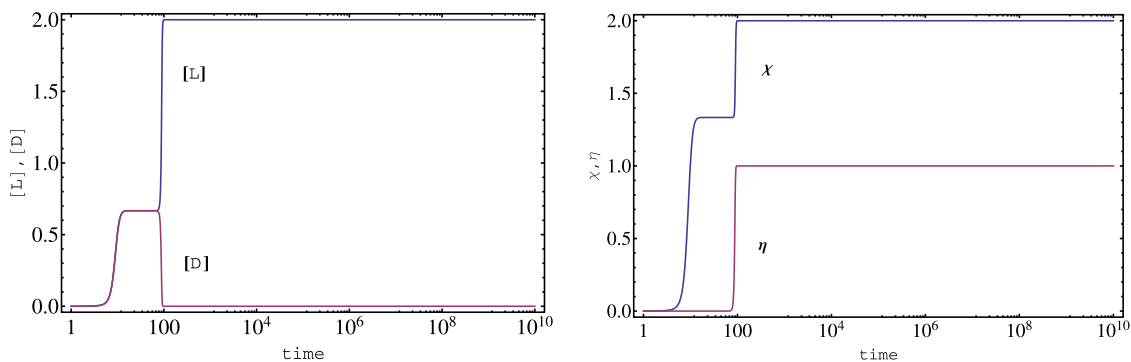


Figure 2. Chiral symmetry breaking in the open system (6)–(8). Temporal behavior (logarithmic scale) of the individual enantiomer concentrations $[L]$ and $[D]$ (left), and the chiral polarization η and total enantiomer concentration χ (right). Initial concentrations: $[L]_0 = (1 \times 10^{-6} + 1 \times 10^{-15})M$, $[D]_0 = 1 \times 10^{-6}M$ ($\eta_0 = 5 \times 10^{-8}\%$) and $[A] = 1M$. Rate constants: $k_1 = 10^{-4}s^{-1}$, $k_{-1} = 10^{-6}s^{-1}$, $k_2 = 1s^{-1}M^{-1}$, $k_{-2} = 0.5s^{-1}M^{-1}$ and $k_5 = 1s^{-1}M^{-1}$. These values correspond to $g = 0.5$ and $u = 10^{-4}$. In figures of simulations, we always display original variables $[L]$, $[D]$, $[LD]$, etc. as function of time t .

increase, not even temporarily. Notice the time scales for χ and η are of the same order. On the other hand, if $g < g_{crit}$, then we have the situation depicted on the right hand side of the figure. Here the same initial point moves towards the vertical nullcline and up towards R , but once past the locally horizontal black curve, is attracted to one of the two chiral fixed points where it stays forever, provided the system is maintained out of equilibrium. The chiral symmetry is permanently broken, and there is no excursion such as we have defined it.

3. Semi-open system

3.1. Rate equations

To elucidate the temporal evolution of χ and η for a more general setting, we do not remove the heterodimer from the system now allow the back reaction to chiral monomers, and we keep $[A]$ constant. There is an implicit inflow as a consequence of constant $[A]$, but no outflow, so we denote this case ‘semi-open’. While there is no mass balance the system can still exhibit temporary SMSB.

The corresponding rate equations are

$$\frac{d}{dt}[L] = k_1[A] + (k_2[A] - k_{-1})[L] - k_{-2}[L]^2 - k_5[L][D] + k_{-5}[LD], \quad (23)$$

$$\frac{d}{dt}[D] = k_1[A] + (k_2[A] - k_{-1})[D] - k_{-2}[D]^2 - k_5[D][L] + k_{-5}[LD], \quad (24)$$

$$\frac{d}{dt}[LD] = k_5[L][D] - k_{-5}[LD]. \quad (25)$$

After performing the same rescaling as in the open case, we arrive at

$$\frac{d\chi}{d\tau} = 2u + \chi - \frac{1}{2}(g+1)\chi^2 - \frac{1}{2}(g-1)\chi^2\eta^2 + 2rP, \quad (26)$$

$$\frac{d\eta}{d\tau} = \eta(1 - g\chi) - \frac{\eta}{\chi} \left(\frac{d\chi}{d\tau} \right), \quad (27)$$

$$\frac{dP}{d\tau} = \frac{1}{4}\chi^2(1 - \eta^2) - rP. \quad (28)$$

The dimensionless parameters appearing here are:

$$u = \frac{k_1 k_5 [A]}{(k_2 [A] - k_{-1})^2}, \quad g = \frac{k_{-2}}{k_5}, \quad r = \frac{k_{-5}}{(k_2 [A] - k_{-1})}. \quad (29)$$

The system is described by three equations, Eqs. (26)–(28) which do not decouple.

3.2. Phase plane and linear stability analysis

In order to obtain an *approximate* two-dimensional phase plane representation of the system, we will invoke the dynamic steady state approximation for the heterodimer. Such approximations are usually justified when there exists a clear separation of time scales in the problem, thus allowing one to identify rapidly and slowly changing concentrations [13]. Here however, no such time scales are evident, all concentration variables evolve on a similar time scale. Nevertheless, we will see a posteriori that this approximation can be good over a wide range of time scales. We therefore assume that the heterodimer is in a dynamic steady state P_{stat} relative to the chiral monomer concentrations and chiral polarization:

$$P_{stat} \approx \frac{\chi^2}{4r} (1 - \eta^2). \quad (30)$$

Substituting this P_{stat} into Eqs. (26) and (27) leads to the differential equations

$$\frac{d\chi}{d\tau} = 2u + \chi - \frac{g}{2}\chi^2(1 + \eta^2), \quad (31)$$

$$\frac{d\eta}{d\tau} = \eta \left(-g\chi - \frac{2u}{\chi} + \frac{g}{2}\chi(1 + \eta^2) \right). \quad (32)$$

As above, we study the phase space of the two-dimensional system by means of the nullclines. These are plotted in Figure 3. The condition $d\chi/d\tau = 0$ implies two curves

$$\chi_{\pm} = \frac{1 \pm \sqrt{1 + 4gu(1 + \eta^2)}}{g(1 + \eta^2)}, \quad (33)$$

whereas the condition $d\eta/d\tau = 0$ implies the three curves

$$\eta = 0, \quad \eta_{\pm} = \pm \sqrt{1 + \frac{4u}{g\chi^2}}. \quad (34)$$

3.3. Fixed points and stability

We solve Eqs. (26)–(28) looking for steady states. To keep the algebra manageable, we also set $u = 0$ as in [9]. There are four solutions, namely, the empty O solution, the racemic R or the two mirror-symmetric chiral Q_{\pm} solutions:

$$O = (P = 0, \chi = 0, y = 0), \quad (35)$$

$$R = \left(P = \frac{1}{g^2 r}, \chi = \frac{2}{g}, y = 0 \right), \quad (36)$$

$$Q_{\pm} = \left(P = 0, \chi = \frac{1}{g}, y = \pm \frac{1}{g} \right). \quad (37)$$

Note that the final heterodimer concentration P is zero in the chiral states Q_{\pm} . Note that the steady state approximation Eq. (30) implies the same result since $|\eta| = 1$. In order to study the linear stability of the four possible homogeneous solutions O, R and Q_{\pm} , we calculate the eigenvalues of the 3×3 Jacobian array M_{open} derived in [9] after deleting the 3rd and 4th rows and columns. The eigenvalues corresponding to these solutions are given by

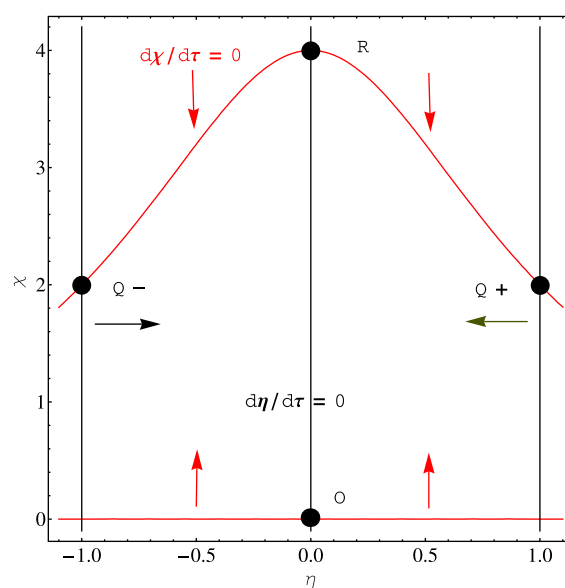


Figure 3. Nullclines for the semi-open case (31) and (32) in the steady state approximation for P . The nullclines (33) and (34) are plotted in red and black, respectively. Red and black arrows indicate the phase-space regions of increasing or decreasing χ and η . The four stationary solutions O, R, Q_{\pm} are indicated by the black dots. These curves are illustrated for $u = 0$ and $g = 0.5$.

$$\lambda_{1,2,3}(O) = (1, 1, -r), \quad (38)$$

$$\lambda_{1,2,3}(R) = \left(-1, -\frac{2 + g(1+r) + \sqrt{4 + g^2(-1+r)^2 + 4g(1+r)}}{2g}, \right. \\ \left. \frac{-2 - g(1+r) + \sqrt{4 + g^2(-1+r)^2 + 4g(1+r)}}{2g} \right), \quad (39)$$

$$\lambda_{1,2,3}(Q_{\pm}) = \left(-1, -\frac{1 + g(-1+r) + \sqrt{1 + 2g(-1+r) + g^2(1+r)^2}}{2g}, \right. \\ \left. \frac{-1 + g(1-r) + \sqrt{(1 + g(-1+r))^2 + 4g^2r}}{2g} \right). \quad (40)$$

As $\lambda_{1,2}(O) > 0$, the empty state is always unstable. An inequality analysis shows that both $\lambda_2(R) < 0$ and $\lambda_3(R) < 0$ for all $r > 0$ and $g > 0$. Since $\lambda_1(R) = -1$ this demonstrates that the racemic state R is always stable. As an independent check, we also verify that $\lambda_3(Q) > 0$ is positive for all $r > 0$ and $g > 0$, so the chiral solutions Q_{\pm} are always unstable. The final outcome will always be the racemic state. There is no stable mirror symmetry broken solution when the heterodimer dissociates back into the chiral monomers. Nevertheless, the system can have temporary chiral excursions.

In Figure 4 we plot the temporal evolution of the L and D chiral monomers starting from an extremely dilute total enantiomer concentration and the very small statistical chiral deviations from the ideal racemic composition. The right hand side of this figure shows the evolution in terms of the quantities χ and η . Note the chiral excursion in η for the time interval between $t \simeq 100$ s and $t \simeq 1000$ s. Finally, we compare the heterodimer concentration from direct numerical simulation with the steady state approximation in Figure 5. To do so we simulate the original concentration variables in Eqs. (23)–(25) and then transform results in terms of χ , η and P . While P_{stat} appears to overestimate $P = [\widetilde{LD}]$, it provides a reasonably good approximation to the actual heterodimer concentration P right after chiral symmetry is broken, at about $t \simeq 100$ s and coincides perfectly after chiral symmetry is recovered, and when χ reaches its asymptotic value, after approximately $t \simeq 10^4$ s).

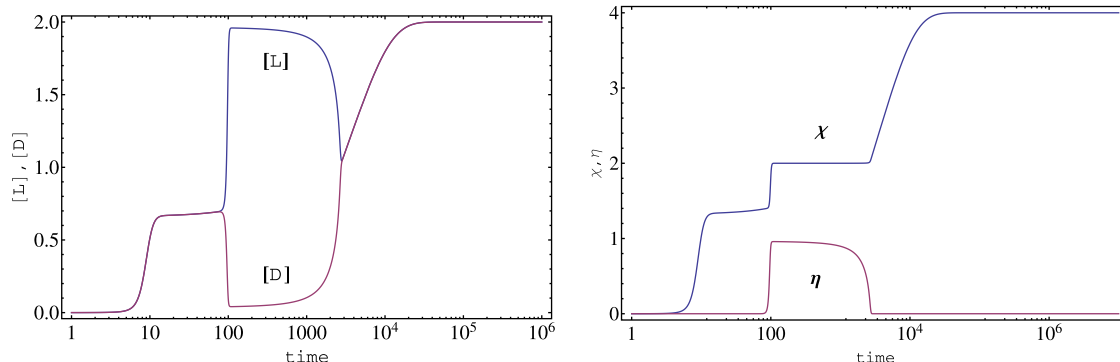


Figure 4. Temporary chiral symmetry breaking and chiral excursions in the semi-open system (23)–(25). Temporal behavior (logarithmic scale) of the individual enantiomer concentrations $[L]$ and $[D]$ (left) and the chiral polarization η and total enantiomer concentration χ (right). Initial concentrations: $[L]_0 = (1 \times 10^{-7} + 1 \times 10^{-15})M$, $[D]_0 = 1 \times 10^{-7}M$ ($\eta_0 = 5 \times 10^{-8}\%$) and $[A] = 1M$. Rate constants: $k_1 = 10^{-3}s^{-1}$, $k_{-1} = 10^{-6}s^{-1}$, $k_2 = 1s^{-1}M^{-1}$, $k_{-2} = 0.5s^{-1}M^{-1}$, $k_5 = 1s^{-1}M^{-1}$ and $k_{-5} = 10^{-3}s^{-1}$. These rate constants imply $g = 0.5$ and $u = 10^{-4}$.

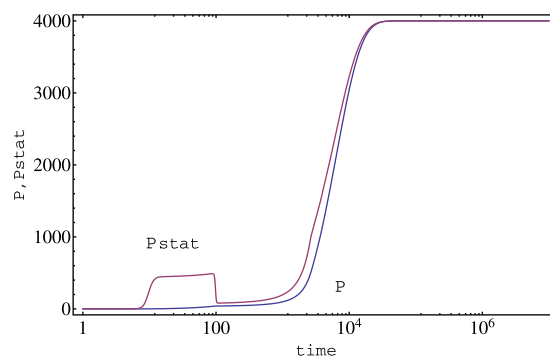


Figure 5. Semi-open system: Comparison of direct numerical solution $P = [\widetilde{LD}]$ and the steady state approximation P_{stat} in Eq. (30) for the heterodimer concentration (after transforming from $[L], [D], [LD]$ to χ, η, P).

4. Closed system

4.1. Rate equations

The rate equations directly follow from (1)–(3):

$$\frac{d}{dt}[L] = k_1[A] + (k_2[A] - k_{-1})[L] - k_{-2}[L]^2 - k_5[L][D] + k_{-5}[LD], \quad (41)$$

$$\frac{d}{dt}[D] = k_1[A] + (k_2[A] - k_{-1})[D] - k_{-2}[D]^2 - k_5[D][L] + k_{-5}[LD], \quad (42)$$

$$\frac{d}{dt}[A] = -2k_1[A] - (k_2[A] - k_{-1})([L] + [D]) + k_{-2}([L]^2 + [D]^2), \quad (43)$$

$$\frac{d}{dt}[LD] = k_5[L][D] - k_{-5}[LD]. \quad (44)$$

There is no flow of material into or out of the system. Since $[A]$ is not constant in this situation, we cannot use it to rescale the time or the concentrations. Instead, we take $\tau = k_1 t$ for the time parameter and $[\tilde{L}] = (k_5/k_1)[L]$, etc. for the dimensionless concentrations. This allows us to express the rate equations in the following dimensionless form:

$$\frac{d}{d\tau}[\tilde{L}] = [\tilde{A}] - u[\tilde{L}] + h[\tilde{A}][\tilde{L}] - g[\tilde{L}]^2 - [\tilde{L}][\tilde{D}] + \rho[\tilde{LD}], \quad (45)$$

$$\frac{d}{d\tau}[\tilde{D}] = [\tilde{A}] - u[\tilde{D}] + h[\tilde{A}][\tilde{D}] - g[\tilde{D}]^2 - [\tilde{D}][\tilde{L}] + \rho[\tilde{LD}], \quad (46)$$

$$\frac{d}{d\tau}[\tilde{LD}] = [\tilde{L}][\tilde{D}] - \rho[\tilde{LD}]. \quad (47)$$

These are subject to the constraint $[\tilde{A}] = [\tilde{C}] - ([\tilde{L}] + [\tilde{D}]) - 2[\tilde{LD}]$, where $[\tilde{C}] = (k_5/k_1)[C]$ and $[C]$ is the total initial concentration, being constant in time. The four parameters appearing here are

$$u = \frac{k_{-1}}{k_1}, \quad g = \frac{k_{-2}}{k_5}, \quad h = \frac{k_2}{k_5}, \quad \rho = \frac{k_{-5}}{k_1}. \quad (48)$$

Changing variables as before to χ, η, P , we arrive at

$$\frac{d\chi}{d\tau} = 2[\tilde{A}] + (h[\tilde{A}] - u)\chi - \frac{1}{2}(g+1)\chi^2 - \frac{1}{2}(g-1)\chi^2\eta^2 + 2\rho P, \quad (49)$$

$$\frac{d\eta}{d\tau} = \eta(h[\tilde{A}] - g\chi - u) - \frac{\eta}{\chi} \left(\frac{d\chi}{d\tau} \right), \quad (50)$$

$$\frac{dP}{d\tau} = \frac{1}{4}\chi^2(1 - \eta^2) - \rho P. \quad (51)$$

In these variables, the constant mass constraint reads $[\tilde{A}] = [\tilde{C}] - \chi - 2P$.

4.2. Phase plane and linear stability analysis

As in the semi-open case, to obtain an approximate two-dimensional phase plane portrait, we assume that the heterodimer is in an approximate steady state P_{stat} and solve Eq. (51) for

$$P_{stat} \approx \frac{\chi^2}{4\rho}(1 - \eta^2). \quad (52)$$

Substituting this back into Eqs. (49) and (50), we obtain

$$\frac{d\chi}{d\tau} = (2 + h\chi)[\tilde{A}] - u\chi - \frac{g}{2}\chi^2(1 + \eta^2), \quad (53)$$

$$\frac{d\eta}{d\tau} = \eta \left(-g\chi - \frac{2[\tilde{A}]}{\chi} + \frac{g}{2}\chi(1 + \eta^2) \right), \quad (54)$$

where

$$[\tilde{A}] = [\tilde{C}] - \chi - \frac{\chi^2}{2\rho}(1 - \eta^2). \quad (55)$$

The nullcline condition $d\chi/d\tau = 0$ leads to an unwieldy cubic equation in χ . More importantly, the nullcline is an *even* function of η , reflecting the underlying Z_2 mirror symmetry. The other condition $d\eta/d\tau = 0$ is straightforward to solve analytically and leads – after discarding the unphysical solution corresponding to negative total enantiomer concentrations – to two curves

$$\eta = 0,$$

$$\chi = \frac{-1 + \sqrt{1 + 2[\tilde{C}](g(1 + \eta^2)/2 + (1 - \eta^2)/\rho - g)}}{g(1 + \eta^2)/2 + (1 - \eta^2)/\rho - g}. \quad (56)$$

The curve $\chi(\eta)$ is an even function of η and $\chi(\eta) \rightarrow \infty$ for $\eta^2 \rightarrow 1$ with a minimum at $\eta = 0$. Thus the nullcline has the form of a narrow fork as depicted in Figure 6.

4.3. Fixed points and stability

A linear stability analysis for the closed Frank model is given in [9]. That analysis was carried out in terms of χ, y and P and does not assume the stationary approximation for the heterodimer. It turns out, even in a model as simple as this one, that keeping $\rho > 0$ is analytically untractable, so we consider $\rho = 0$ in what follows. Then the asymptotic stationary racemic R and chiral solutions Q_{\pm} are given by

$$R = \left(P = \frac{[\tilde{C}]}{2}, \chi = 0, y = 0 \right), \quad (57)$$

$$Q_{\pm} = \left(P = \frac{[\tilde{C}]g + u}{2g}, \chi = -\frac{u}{g}, y = \pm \frac{u}{g} \right). \quad (58)$$

The chiral solutions Q_{\pm} are unphysical for all $u > 0$ since they imply negative total enantiomer concentrations $\chi < 0$. Thus, the only physically acceptable solution is the racemic one R , and this is (at

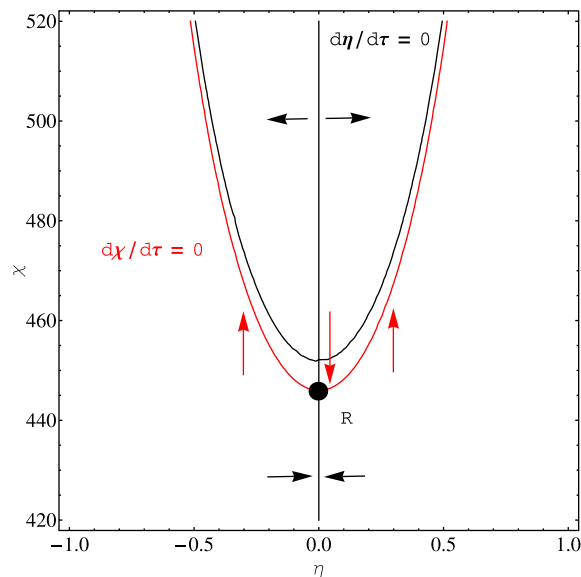


Figure 6. Nullclines for the closed system (53) and (54) in the steady state approximation for P . The η and χ nullclines are plotted as the black and red curve, respectively. These nullclines intersect at the one point (indicated with a dot) which corresponds to the stable racemic solution R with $\eta = 0$ and $\chi > 0$. The flow directions are indicated with arrows. This diagram corresponds to the parameter values $g = 0.05$, $u = 0.01$, $h = 0.1$ and $\rho = 1.0$ and $[\tilde{C}] = 10^5$.

least marginally) stable, the corresponding eigenvalue was calculated to be [9]

$$\lambda_{1,2,3}(R) = (0, -2 - u, -u). \quad (59)$$

In the limit $\rho = 0$, the substrate is consumed and all the matter ends up finally as pure heterodimer. Finally, note that $\lim_{u \rightarrow 0} Q_{\pm} = R$: the unphysical chiral solutions merge to the racemic one when $k_{-1} = 0$. For reversible heterodimer, the matter in the racemic state $\eta = 0$ is distributed between the chiral monomers and the heterodimer: $P = ([\tilde{C}] - \chi)/2$ and $\chi > 0$ in keeping with the law of mass action. The single intersection R displayed in Figure 6 indicates that the racemic state is the only possible solution, in qualitative agreement with the stability analysis.

Example of a chiral excursion in a closed system is provided in Figure 7. Once again, as for the semi-open situation, the scheme is capable of amplifying a tiny initial chiral excess to practically 100%, followed by final approach to the racemic state. The steady state approximation for the heterodimer is rather poor during the early stages (Figure 8), but similar to semi-open case, converges to the true heterodimer concentration after symmetry breaking and restoration.

5. Discussion

In this Letter, we investigated transient mirror symmetry breaking in chiral systems, in particular in the Frank model in settings of open, semi-open, and closed environments. Temporary chiral excursions are observed for closed and semi-open systems and explained through phase space analysis, stability analysis and numerical simulations. Such chiral excursions may be experimentally observed and could be mistaken for a transition to a chiral state. They are in fact a long sought goal for the experimental chemist who could actually fail to see them if not aware of their transitory nature. In open systems by contrast, the racemic state is approached monotonically. Therefore, it is important to understand the processes and constraints responsible for these

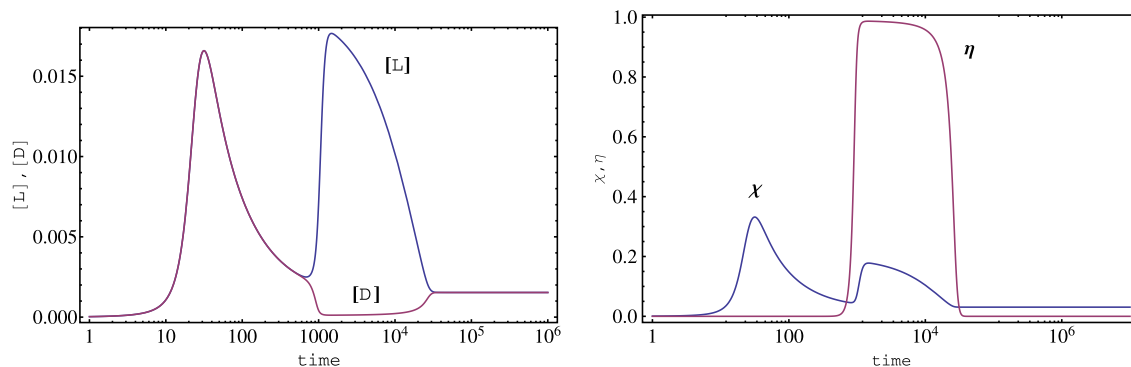


Figure 7. Temporary chiral symmetry breaking and chiral excursions in the closed system (41) and (42). Temporal behavior (logarithmic scale) of the individual enantiomer concentrations $[L]$ and $[D]$ (left) and the chiral polarization η and total enantiomer concentration χ (right). Initial concentrations: $[L]_0 = (1 \times 10^{-7} + 1 \times 10^{-15})M$, $[D]_0 = 1 \times 10^{-7}M$ ($\eta_0 = 5 \times 10^{-8}\%$) and $[A]_0 = 1M$. Rate constants: $k_1 = 10^{-4}s^{-1}$, $k_{-1} = 10^{-6}s^{-1}$, $k_2 = 1s^{-1}M^{-1}$, $k_{-2} = 0.5s^{-1}M^{-1}$, $k_5 = 10s^{-1}M^{-1}$ and $k_{-5} = 10^{-4}s^{-1}$.

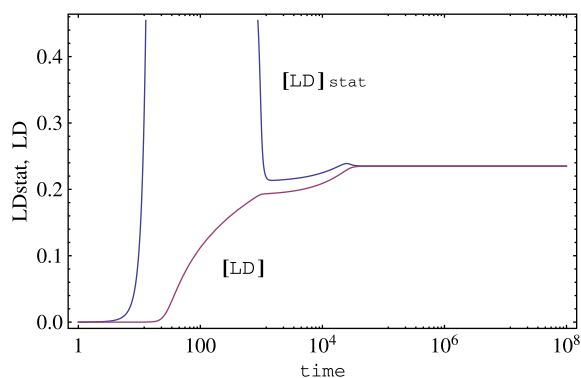


Figure 8. Comparison of direct numerical solution $[LD]$ and the steady state approximation $[LD]_{stat} = (k_5/k_{-5})[L][D]$ for the heterodimer concentration. The steady state approximation peaks to a maximum value of about 25 in this example.

outcomes. This Letter has focused on the effects that the in- and outflow of matter has on these phenomena. The open nature of the Frank model can be arranged *experimentally* with an incoming flow of achiral precursor A and an outflow of the product heterodimer LD, to conserve mass balance. Mathematically, we can model the inflow by assuming a constant concentration of $[A]$ and the outflow by a term representing the rate at which LD leaves the system. In our open model we assumed no dissociation of LD back into chiral monomers. From the point of view of achieving permanent SMSB, there is then actually no need to remove LD from the system, but we retain this outflow since it is needed to ensure stationary fixed points for all the chemical concentrations. For the semi-open case, LD is not removed and we allow for its dissociation into chiral monomers. There is no mass balance but temporary symmetry breaking can arise. Finally, in the closed system there is neither inflow nor outflow, total mass is conserved and temporary symmetry breaking can occur.

A recent kinetic analysis of the Frank model in closed systems applied to the Soai reaction [5] indicates that in actual chemical scenarios, reaction networks that exhibit SMSB are extremely sensitive to chiral inductions due to the presence of inherent tiny initial enantiomeric excesses [4]. This amplification feature is also operative in much more involved reaction networks such as chiral polymerization [6]. When the system is subject to a very small perturbation about an extremely dilute racemic state, the initial chiral fluctuation does not immediately decay, but becomes amplified and drives the system along a long-lived chiral excursion in phase space before final and inevitable approach to the stable racemic

solution. Mauksch and Tsogoeva have also previously indicated that chirality could appear as the result of a temporary asymmetric amplification [14,15].

Excursions in phase space as studied in this work are superficially reminiscent of excitable systems as studied in dynamical systems [13,16,17]. But there are important differences. First of all, the excursions reported in chiral systems are not easily visualized in the chiral monomer concentrations themselves, but are strikingly manifested by the chiral polarization or enantiomeric excess. Secondly, the total enantiomer and heterodimer concentrations do increase with time, so that the complete phase-space trajectory does not follow a closed path: there is no return to the initial state. The chiral excursion is a one way trip, not a round trip as in an excitable system. Third, whereas excursions are traditionally studied for open excitable systems [13,16,17], chiral excursions are observed here only for closed or at most semi-open systems, but not for open systems.

The original impetus for considering phase-space descriptions of the Frank model comes not only from the chiral excursions reported in [4] and [6] but also by the recent report of damped chiral oscillations detected numerically in a model of chiral polymerization in closed systems [6]. Absolute asymmetric synthesis is achieved in the latter scheme, accompanied by long duration chiral excursions in the enantiomeric excesses for all the homopolymer chains formed, analogous to the much simpler Frank model. But unlike the latter, strong enantiomeric inhibition converts these excursions into long period damped chiral oscillations in the enantiomeric excesses associated with the longest homochiral polymer chains formed. Moreover, short period sustained chiral oscillations have been observed numerically in a recycled Frank model open to energy flow, for large values of the inhibition [18]. This oscillatory behavior poses an additional problem for the origin of biological homochirality, since any memory of the sign of the initial fluctuation is further erased by subsequent oscillations thus adding a further element of uncertainty to the overall problem. Chemical oscillations have been traditionally studied in conjunction with excitability. Although the latter concept is not directly applicable to models exhibiting SMSB, it remains to be seen if the techniques used to study oscillations can be applied profitably to reaction schemes that lead to chiral oscillations.

Acknowledgments

DH acknowledges the Grant AYA2009-13920-C02-01 from the Ministerio de Ciencia e Innovación (Spain) and forms part of the ESF COST Action CM07030: *Systems Chemistry*. MS acknowledges support from the Spanish MICINN through project FIS2008-05273

and from the Comunidad Autónoma de Madrid, project MODELICO (S2009/ESP-1691). CB acknowledges a Calvo-Rodés scholarship from the Instituto Nacional de Técnica Aeroespacial (INTA).

References

- [1] F.C. Frank, *Biochim. et Biophys. Acta* 11 (1953) 459.
- [2] S.F. Mason, *Chemical Evolution*, Oxford University Press, Oxford, 1991.
- [3] R. Plasson, D.K. Kondepudi, H. Bersini, A. Commeyras, K. Asakura, *Chirality* 19 (2007) 589.
- [4] J. Crusats, D. Hochberg, A. Moyano, J.M. Ribó, *ChemPhysChem* 10 (2009) 2123.
- [5] K. Soai, K.T. Shibata, H. Morioka, K. Choji, *Nature* 378 (1995) 767.
- [6] C. Blanco, D. Hochberg, *Phys. Chem. Chem. Phys.* 13 (2011) 839.
- [7] I. Weissbuch, R.A. Illos, G. Bolbach, M. Lahav, *Acc. Chem. Res.* 42 (2009) 1128.
- [8] T. Hitz, P.L. Luisi, *Helv. Chim. Acta.* 86 (2003) 1423.
- [9] J.M. Ribó, D. Hochberg, *Phys. Lett. A* 373 (2008) 111.
- [10] L.L. Morozov, V.V. Kuzmin, V.I. Goldanskii, *Orig. Life* 13 (1983) 119.
- [11] D.K. Kondepudi, G.W. Nelson, *Physica A* 125 (1984) 465.
- [12] V.A. Avetisov, V.V. Kuzmin, S.A. Anikin, *Chem. Phys.* 112 (1987) 179.
- [13] S.K. Scott, *Oscillations, Waves, and Chaos in Chemical Kinetics*, Oxford University Press, Oxford, 1994.
- [14] M. Mauksch, S.B. Tsogoeva, S.-W. Wei, I.M. Martinova, *Chirality* 19 (2007) 816.
- [15] M. Mauksch, S.B. Tsogoeva, *ChemPhysChem* 9 (2008) 2359.
- [16] E. Meron, *Phys. Rep.* 218 (1992) 1.
- [17] Albert Goldbeter, *Biochemical Oscillations and Cellular Rhythms*, Cambridge University Press, Cambridge, 1996.
- [18] R. Plasson, H. Bersini, A. Commeyras, *Proc. Natl. Acad. Sci. USA.* 101 (2004) 16733.

DOI: 10.1002/cphc.201300350

Spontaneous Emergence of Chirality in the Limited Enantioselectivity Model: Autocatalytic Cycle Driven by an External Reagent

Celia Blanco,^[a] Joaquim Crusats,^[b, c] Zoubir El-Hachemi,^[b, c] Albert Moyano,^[b] David Hochberg,^{*[a]} and Josep M. Ribó^{*[b, c]}

The model of limited enantioselectivity (LES) in closed systems, and under experimental conditions able to achieve chemical equilibrium, can give rise to neither spontaneous mirror symmetry breaking (SMSB) nor kinetic chiral amplifications. However, it has been recently shown that it is able to lead to SMSB, as a stationary final state, in thermodynamic scenarios involving nonuniform temperature distributions and for compartmentalized separation between the two autocatalytic reactions. Herein, it is demonstrated how SMSB may occur in LES in a cyclic network with uniform temperature distribution if the reverse reaction of the nonenantioselective autocatalysis,

which gives limited inhibition on the racemic mixture, is driven by an external reagent, that is, in conditions that keep the system out of chemical equilibrium. The exact stability analysis of the racemic and chiral final outcomes and the study of the reaction parameters leading to SMSB are resolved analytically. Numerical simulations, using chemical kinetics equations, show that SMSB may occur for chemically reasonable parameters. Numerical simulations on SMSB are also presented for speculative, but reasonable, scenarios implying reactions common in amino acid chemistry.

1. Introduction

Most applied chemists view the theoretical models of spontaneous mirror symmetry breaking (SMSB) in chemical transformations as mere academic research efforts, because of the absence of actual experimental models to study. This despite the recent reports on absolute asymmetric synthesis^[1] and on the spontaneous deracemization of racemic mixtures of enantiopure crystals.^[2] The significance of these reports is not easily grasped by synthetic organic chemists, not only because they deal with the rather uncommon phenomenon of enantioselective autocatalysis, but also as a consequence of the lack of reliable models for explaining SMSB in chemical processes.^[3,4] The interest in finding the correct chemical explanations of these results is not only related to the phenomenon of biological homochirality, but also to the prospects of achieving absolute asymmetric synthesis in applied chemistry.

Enantioselective autocatalysis is the key reaction that may lead to SMSB on coupling to reactions that amplify the enan-


tiomeric excess (*ee*) generated by statistical fluctuations about the ideal racemic composition, or induced by small chiral polarizations. For certain reaction parameters and thermodynamic constraints, such a system as a whole follows a bifurcation scenario in which the racemic final state is metastable and the more stable final state is chiral.^[5] Clearly, for a system in conditions that allow a chemical thermodynamic equilibrium to be reached (i.e. a closed system with a uniform distribution of matter, temperature, and energy), the final outcome must be the racemic state. However, in a system with a nonuniform energy distribution (e.g. energy absorption by only some of the species of the system, or open to matter exchange with the surroundings), the final stable stationary state may be chiral. A classical example of this is called the Frank model of SMSB in open systems.^[6] Furthermore, a Frank-like reaction network (see below) in a closed system (Scheme 1, left-hand side) with a uniform temperature distribution can lead to chiral excursions, that is, to kinetically controlled temporary chiral windows, which in the case of exergonic transformations are resilient to racemization that occurs on a timescale much longer than that required for the fast conversion of the reactants to products.^[7] This chiral window can be so wide that, from the point of view of organic synthesis, the reaction is an absolute asymmetric synthesis, such as has been demonstrated in the case of the Soai reaction.^[1]

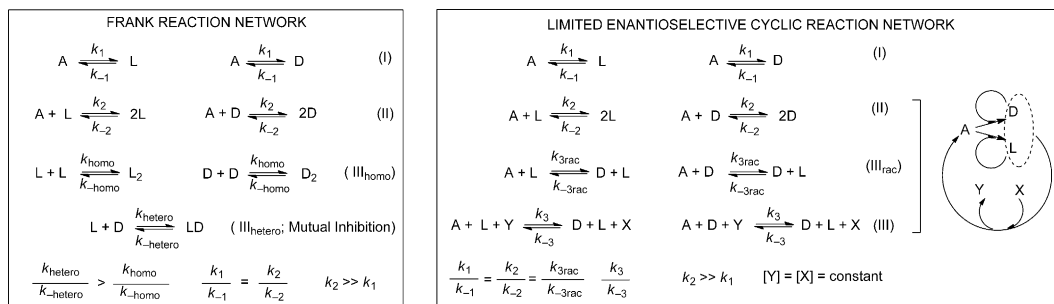
The limited enantioselectivity (LES) model^[8] (Scheme 1, right-hand side) is composed of an enantioselective autocatalytic reaction coupled to the corresponding nonenantioselective autocatalysis, and was proposed originally as a SMSB alternative to the Frank model.^[6] The Frank reaction network is composed of

[a] C. Blanco, Dr. D. Hochberg
Centro de Astrobiología (CSIC-INTA)
Ctra. Ajalvir Km 4, 28850 Torrejón de Ardoz, Madrid (Spain)
E-mail: hochbergd@cab.inta-csic.es

[b] Dr. J. Crusats, Dr. Z. El-Hachemi, Prof. A. Moyano, Prof. J. M. Ribó
Departament de Química Orgànica
Universitat de Barcelona
C. Martí I Franquès 1, 08028 Barcelona, Catalonia (Spain)
E-mail: jmribo@ub.edu

[c] Dr. J. Crusats, Dr. Z. El-Hachemi, Prof. J. M. Ribó
Institute of Cosmos Science (IEEC-UB)
C. Martí I Franquès 1, 08028 Barcelona, Catalonia (Spain)

 Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cphc.201300350>.



Scheme 1. Frank^[6] and limited enantioselective^[8] cyclic reaction networks. For a description of the kinetically controlled SMSB in a closed system of this Frank model, see ref. [7d]. The LES network cannot lead to absolute asymmetric synthesis, due to the microreversibility principle, when it is composed of the three reactions (I) + (II) + (III_{rac}) (see ref. [13]). However, it leads to SMSB in a compartmentalized system at different temperatures.^[15] The cyclic LES network studied herein is that composed of the reactions (I) + (II) + (III) in a uniform temperature distribution and at constant concentrations of the reactants Y and X, that is, an open system with Y and X matter exchange. A is an achiral compound and D and L the enantiomer pair of a chiral compound.

the coupling between the enantioselective autocatalysis (II) and a reaction of mutual inhibition between the enantiomeric reaction compounds/catalysts (III_{hetero}). In both models the enantioselective autocatalysis (II) is coupled to a reaction leading to the decrease of the racemic composition of chiral catalysts.^[9] This leads to an increase in the *ee* that is strongly amplified in the basic reaction (II).^[10] The amplification is so strong that statistical fluctuations about the ideal racemic mixture^[10] may be amplified up to detectable *ee* values. The significant difference between the Frank and LES models is that in the Frank model the reaction (III_{hetero}) must be more exergonic than the (III_{homo}) one.^[7d] Notice that this happens for many but not all chiral compounds,^[11] whereas the LES reaction network does not require such a mutual inhibition between the enantiomers.

The LES model has had a controversial reputation because at times the thermodynamic constraints imposed by the principle of microreversibility have not been considered,^[7d,12] that is, the simulated SMSB obtained in those cases cannot represent a thermodynamically possible process. In fact, it has been shown^[13] that LES in a closed system does not lead to SMSB, nor can it lead to a temporary and resistant amplification^[14] of the initial *ee*, such as those that can be obtained in Frank-like networks for relatively high exergonic reactions.^[7d] However, recent reports demonstrate that LES reaction networks may lead to SMSB in a scenario in which the enantio- and nonenantioselective autocatalyses are each individually compartmentalized within different temperature regions of the system.^[15] Taking into account the dependence of the equilibrium constant and rate constants on the temperature, a very high temperature gradient between compartments is necessary for such a SMSB. This determines that the medium can only remain in the fluid state in all regions of the system in the case of enormous pressures, such as those present in deep ocean hydrothermal plumes, and this has been suggested as a plausible prebiotic scenario for SMSB.^[15a]

Herein, we discuss the SMSB in the LES model with a uniform temperature distribution, but with the nonenantioselective autocatalysis driven by a constant concentration of external reactants (X and Y in reaction (III) in Scheme 1, right-hand side), that is, an open system with X and Y matter exchange with

the surroundings. Notice that, regarding the substrate and the final products, the reaction network is a cyclic one with permanent consumption and production of X and Y, respectively, or of Y and X depending on the reaction flow direction in the cycle.

We show how this type of system can achieve SMSB for a wide range of reaction parameters. In Section 3 we perform a stability analysis and discuss the conditions for SMSB of the minimal LES model (I) + (II) + (III) in a uniform temperature distribution. However, reaction (III) in Scheme 1 is difficult to imagine as a chemically reasonable reaction because of its high molecularity. Therefore, in Section 4 we discuss the conditions for an applied chemistry scenario through the simulation of a system with reasonable chemical reactions in a cycle composed of the amino acid Strecker synthesis and the amino acid Strecker degradation.

2. Numerical Simulations

Simulations were performed by numerical integration of the differential rate equations according to rate-equation theory as applied in chemical kinetics (mean field assumption). The concentration units are mol L⁻¹ and the different rate constants have the appropriate units to yield rate values in units of mol s⁻¹. Numerical integration was performed with the Mathematica program package. The results were monitored and verified to ensure that the total system mass remained constant in time. For a set of parameters corresponding to the system at, or very near to, the bifurcation point, the numerical integration is highly sensitive to minute differences between the reaction parameters, so that the inherent numerical noise of the calculations suffices to bifurcate the system towards a chiral outcome, or is made insensitive to SMSB. In our simulations we suppressed this computational noise, which arises from round-off errors, by setting a high numerical precision of the input parameters (100 significant decimal digits) and exact number representation of the reaction rates and the initial concentration values (e.g. "1+1×10⁻²⁰" instead of "1.01" or "1.+1×10⁻¹⁰" or "1+1.×10⁻²⁰"). Integration methods of "StiffnessSwitching" and "WorkingPrecision" of up to 50 were used in the present calculations. An example of the Mathematica input for

one of the examples reported in Section 3 is provided in the Supporting Information. The fluctuations of chirality able to convert the racemic output to a chiral one were simulated by using an initial *ee* of products/catalysts lower than that expected from the statistical fluctuations about the ideal racemic composition, that is, an initial *ee* (%) < 67.43 × (N^{-0.5}), in which *N* is the number of chiral molecules.^[10] By using this procedure the SMSB was detected for initial *ee* values much lower than those of the statistical fluctuations from the ideal racemic composition and in the absence of any chiral polarization. In conditions of an initial ideal racemic composition of chiral compounds, or in their initial absence, no SMSB takes place, that is, the composition of the virtual metastable racemic state could be determined. The numerical integration was run between 0 and 1 × 10²⁰ s. This limit of time, three orders of magnitude larger than the age of the universe, was used to compare the simulated numerical values of the *ee* and the final concentrations of chiral compounds with those obtained from the stability analysis at *t* → ∞ [see Eqs. (4)–(6) below] and those obtained from the asymptotic analytic stability analysis. In this respect, the values of the final concentrations [A], [D], and [L] calculated through the stability analysis [see Eqs. (4)–(6) below] and those obtained by the numerical integration agree at the level of the numerical errors in the calculated precision.

3. Results

3.1. Analysis of the Conditions to Yield a Final Chiral Stationary State in the Network (I) + (II) + (III)

We can take over the stability analysis previously carried out for LES [(I) + (II) + (III)_{rac}] (see Scheme 1, right-hand side) in closed systems by formally substituting *k*_{3→}[Y]*k*₃ and *k*_{-3→}[X]*k*₋₃ in the relevant expressions for the static solutions and eigenvalues that appear in Section 5.3 in ref. [13]. For this purpose, the following dimensionless parameters are defined as follows, taking into account that [Y] and [X] are constant concentration values [Eq. (1)]:

$$\begin{aligned} u &= k_{-1}/k_1; \quad g = k_{-2}/([X]k_{-3}); \\ h &= k_2/([X]k_{-3}); \quad r = ([Y]k_3)/([X]k_{-3}) \end{aligned} \quad (1)$$

Note that the principle of microreversibility requires [Eq. (2)]:

$$k_1/k_{-1} = k_2/k_{-2} \quad (2)$$

and therefore [Eq. (3)]:

$$g/h = u \quad (3)$$

Note that in the case of [Y] = [X], *r* is the thermodynamic equilibrium constant of the transformation (III) and in the case of unequal concentrations [Y] ≠ [X], it is the corresponding effective equilibrium ratio.

The complete set of reaction parameters defining the network are the six rate constants *k*_{±*i*} (*i* = 1, 2, 3) subject to the

constraint Equation (2), the initial concentrations [A]₀, [L]₀, and [D]₀, and the constants [Y] and [X], 11 in all. The constriction of Equation (2) reduces the total number of independent parameters defining the system to 10. The dimensionless parameters of Equation (1) are determined in terms of these, and the constraint of Equation (3) should hold automatically. Therefore, the final states depend only on the four dimensionless parameters in Equation (1), subject to the one constraint Equation (3), the total concentration *C*, and the conversion factor [X]*k*₋₃/*k*₁, that is, a total of five parameters.

The dimensionless concentrations *A*, *L*, and *D* are related to the dimensionfull concentrations by, for example, *A* = ([X]*k*₋₃/*k*₁)[A]. We remark that the dimensionless total system concentration is defined by *C* = *A* + *L* + *D*.

3.1.1. Exact Solutions for the Final Stationary Racemic (R) and Stationary Chiral (Q_±) Solutions

We generalize the results in ref. [13] as follows in that we include the exact full dependence on *u* (i.e. for *u* > 0). The temporal asymptotic *R* and Q_± stationary solutions are given by [Eq. (4)]:

$$R \equiv \left(y = 0, \chi = \frac{-2 + C(h+r) - u}{1 + g + 2(h+r)} + \frac{\sqrt{4C(1 + g + 2(h+r)) + (2 - C(h+r) + u)^2}}{1 + g + 2(h+r)} \right) \quad (4)$$

and [Eq. (5)]:

$$Q_{\pm} = \left(y = \pm \left(\frac{-4Cg(g+h-r) - C^2(h-r)((g-1)h+r+3gr)}{(g-1)(g+h-r)^2} + \frac{+2Cu(h(g-2r-1) + r(g+2r+1))}{(g-1)(g+h-r)^2} + \frac{u(-g(u+4) - 4h + 4r(u+1) + u)}{(g-1)(g+h-r)^2} \right)^{1/2}, \chi = \frac{C(h-r) - u}{g+h-r} \right) \quad (5)$$

in which *y* = *L* - *D* and *χ* = *L* + *D*. For conditions leading to the chiral solution we can evaluate the final *ee* value from using the expressions *y* and *χ* given in Q_±. Thus [Eq. (6)]:

$$ee \text{ (\%)} = (y/\chi) \times 100 \quad (6)$$

To compare *χ* with the value of the dimensionfull concentration [L] + [D] the conversion factor [X]*k*₋₃/*k*₁ must be employed.

3.1.2. Exact Eigenvalues for Racemic State (R) and Stability Criteria

It is sufficient to establish the stability criteria of the final state in terms of the racemic solution *R* in Equation (4) alone. The stability of time asymptotic racemic solution is governed by the algebraic signs of the associated pair of eigenvalues [Eq. (7)]:

$$\lambda_1(R) = -\sqrt{4C(g+2h+2r+1) + (-C(h+r) + u+2)^2}$$

$$\lambda_2(R) = C(h-r) - u - \frac{(g+h-r)}{1+g+2h+2r}(-u+2) + C(h+r) + \sqrt{C^2(h+r)^2 + 4C(1+g+h+r) - 2Cu(h+r) + (u+2)^2}$$
(7)

Stability in state R requires both these eigenvalues to be negative; $\lambda_{1,2}(R) < 0$. The first expression $\lambda_1(R) < 0$ is always negative, whereas the second $\lambda_2(R)$ can be positive or negative depending on the values of the parameters C , g , h , r , and $u=g/h$. The stability of the state R is determined by the algebraic sign of $\lambda_2(R)$ alone. So, we find $\lambda_2(R) < 0$ if and only if $g > g_{\text{crit}}$ for which [Eq. (8)]:

$$g_{\text{crit}}(u) = -\frac{1}{8C}(4C(h-r) + C^2h^2 + 2C^2hr - 3C^2r^2 + 4u - 2Cu(h+r) + u^2 + (C(r-h) + u)(C^2(h+3r)^2 + 8C(2+h+3r) - 2C(h+3r)u + (4+u)^2)^{1/2})$$
(8)

We find the sign of g_{crit} to be real and positive if and only if $h > r$ and [Eq. (9)]:

$$C > C_{\text{min}} = \frac{u+2ru}{h-r} + 2\sqrt{\frac{u(h-r) + r^2u^2}{(h-r)^2}}$$
(9)

Thus, if either of these two conditions is not satisfied (either $h \leq r$ or $C \leq C_{\text{min}}$), then $g_{\text{crit}} < 0$ and because g is always positive, there can be no values of g satisfying $g < g_{\text{crit}}$ and consequently the only stable state will be R the racemic one.

To summarize [Eq. (10)]:

$$h \leq r \text{ OR } C \leq C_{\text{min}} : g_{\text{crit}} < 0 \rightarrow \text{Racemic } (R)$$
(10)

On the other hand, $h > r$ AND $C > C_{\text{min}}$ imply that $g_{\text{crit}}(u) > 0$. In which case we have two alternatives that depend on g : namely, either a final chiral outcome Q_{\pm} or a final racemic one R [Eqs. (11) and (12)]:

$$g < g_{\text{crit}} \rightarrow Q_{\pm}$$
(11)

$$g > g_{\text{crit}} \rightarrow R$$
(12)

Finally, in the limit of large C , and using Equation (3), the critical value becomes [Eq. (13)]:

$$g_{\text{crit}} \rightarrow \frac{1-\frac{r}{h}}{1+\frac{3r}{h}} + \frac{uh}{8}\left(1+\frac{7r}{h}\right) = \frac{1-\frac{r}{h}}{1+\frac{3r}{h}} + \frac{g}{8}\left(1+\frac{7r}{h}\right)$$
(13)

All simulations performed by numerical integration show, at long reaction times, the achievement of either a racemic or a chiral final state in accord with the predictions summarized in Equations (10) and (11) or (12), respectively. The numerical simulations (see below) reveal a clear picture of the chiral final state in complete accord with the theoretical bifurcation scenario depicted by Equations (4) and (5). After the bifurcation point giving the transition from the racemic state to the chiral state, the racemic solution is a metastable state with respect to the more stable degenerate chiral solutions. The higher stability is characterized by a higher total chiral matter of the chiral solutions compared to the metastable racemic state [the χ values of Eqs. (4) and (5)]. In contrast, in the branch of the racemic solution its χ value is higher than that of the chiral solutions. Furthermore, the calculated χ values for all the racemic and metastable chiral states [Eq. (4)] and of the chiral solutions [Eq. (5)] agree at the level of the numerical error with the values obtained in the numerical simulations.

3.1.3. Surface Representation of the Racemic and Chiral Regions

The intersection of the g and g_{crit} hypersurfaces defines the regions leading to racemic or chiral solutions. Figure 1 shows 3D representations of the g surface as a function of u and h values, and the g_{crit} surfaces correspond to varying u and h for different fixed r and C values. The intersection between the g and g_{crit} surfaces divides the racemic and chiral regions of the final-state solutions. The red and green points plotted in Figure 1 correspond to the parameter values of the chemical kinetics simulations displayed in Figures 2 and 3.

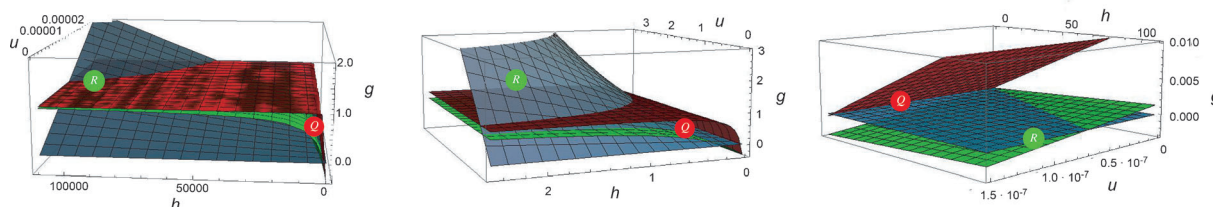


Figure 1. Surfaces of g (blue/gray) and g_{crit} at fixed r and C values. Left: $r=20$ and $C=5$ (red); $r=1000$ and $C=0.1$ (green). Middle: $r=0.02$ and $C=50000$ (red); $r=0.2$ and $C=5000$ (green). Right: $r=0.1$ and $C=5 \times 10^{-4}$ (red); $r=1$ and $C=5 \times 10^{-5}$ (green). The graphics correspond to the r , h , and C values of Figures 2 and 3; the points on the blue g surface correspond to the examples leading to racemic or chiral final states depending on whether the g value lies respectively above or below the g_{crit} value given by the corresponding g_{crit} surface.

$$[A]_0 = 0.1 \text{ mol L}^{-1}; [D]_0 = 1 \cdot 10^{-10} \text{ mol L}^{-1}; [L]_0 = ([D]_0 + 1 \cdot 10^{-20}) \text{ mol L}^{-1}; [Y] = [X] = 0.01 \text{ mol L}^{-1}$$

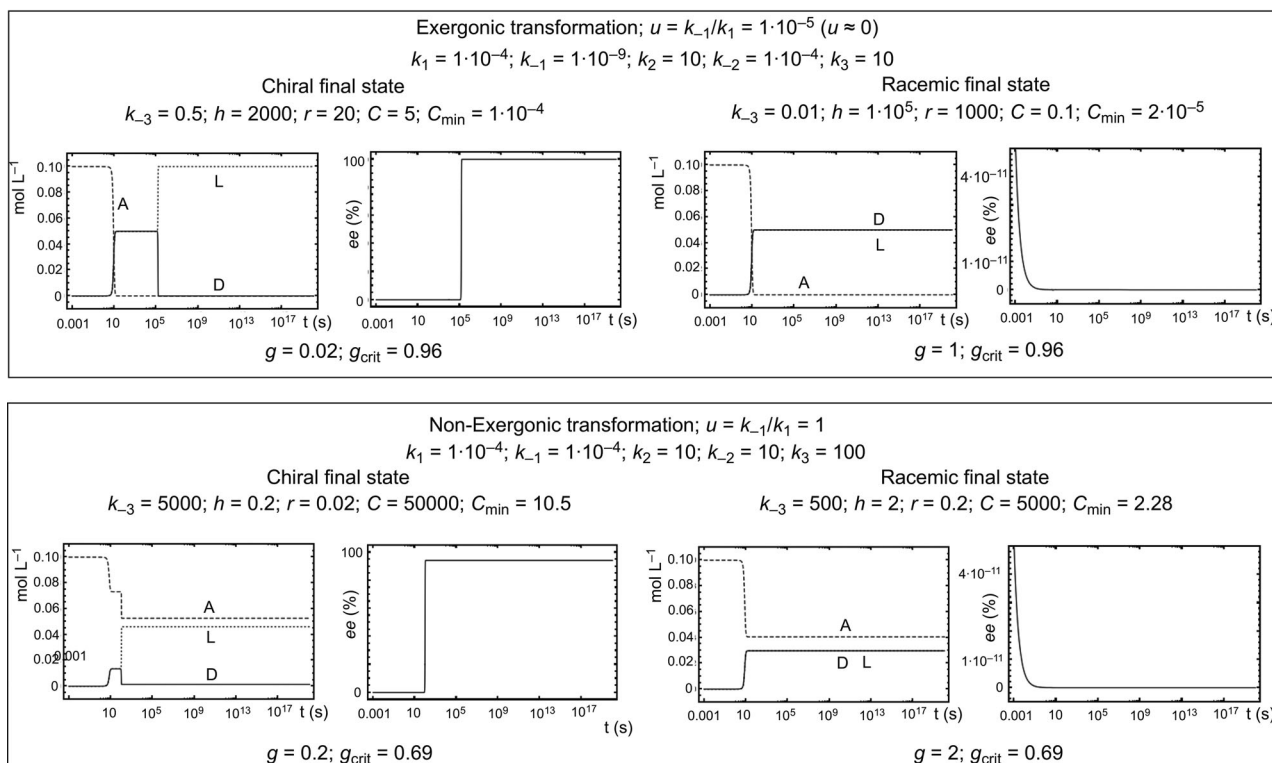


Figure 2. Examples of the time evolution of the LES reaction network [(I) + (II) + (III); see Scheme 1] to achieve the final stable state, either racemic or chiral. The integrations start from an *ee* value below the statistical fluctuations from the ideal racemic composition. The examples correspond to systems showing an enantioselective autocatalysis (II) much faster than the nonautocatalytic reaction (I) ($k_2 \gg k_1$), for an exergonic reaction ($u \approx 0$, top row) and a nonexergonic reaction ($u = 1$, bottom row). The final state obeys the criteria of Equations (11) and (12), that is, when $h > r$ or $C > C_{\min}$, and the final concentrations agree with those calculated from the asymptotic solutions ($t \rightarrow \infty$) of Section 3.1.

$$[A]_0 = 0.5 \text{ mol L}^{-1}; [D]_0 = 1 \cdot 10^{-10} \text{ mol L}^{-1}; [L]_0 = ([D]_0 + 1 \cdot 10^{-20}) \text{ mol L}^{-1}; [Y] = [X] = 0.01 \text{ mol L}^{-1}$$

Exergonic transformation; $u = k_{-1}/k_1 = 1 \cdot 10^{-7}$ ($u \approx 0$)

$$k_1 = 1 \cdot 10^2; k_{-1} = 1 \cdot 10^{-5}; k_2 = 1; k_{-2} = 1 \cdot 10^{-7}; k_3 = 1$$

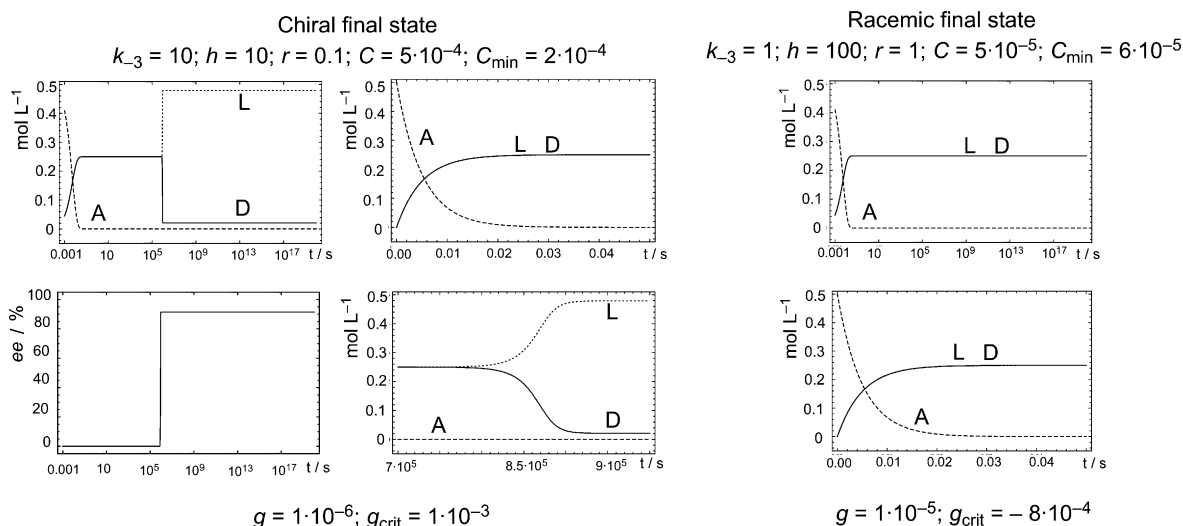


Figure 3. Examples of the time evolution of the LES reaction network [(I) + (II) + (III); see Scheme 1] to achieve the final stable state, either racemic or chiral. The integrations start from an *ee* value below the statistical fluctuations from the ideal racemic composition. The examples correspond to an exergonic reaction ($u \approx 0$) showing an enantioselective autocatalysis (II) slower than the nonautocatalytic reaction (I). The final state obeys the criteria of Equations (11) and (12), that is, when $h > r$ and $C > C_{\min}$, and the final concentrations agree with those calculated from the asymptotic solutions ($t \rightarrow \infty$) of Section 3.1.

3.2. Examples of Chemical Kinetic Simulations

Figure 2 shows two examples of how, for the same reaction parameters, the change of k_{-3} may lead to chiral or racemic final states.^[16] This simple change of one of the ten reaction parameters leads to the change of three (h , r , and C) of the five dimensionless parameters. The first example corresponds to an exergonic reaction ($k_1/k_{-1}=k_2/k_{-2}=1\times 10^5$; that is, at 300 K a reaction $\Delta G^\circ \approx -28.7 \text{ kJ mol}^{-1}$) and the second example illustrates a nonexergonic transformation ($k_1/k_{-1}=k_2/k_{-2}=1$). Both examples show how the deracemization occurs later than that of the transformation of A to D and L.

The comparison between the two pairs of examples shows how flexible^[16] and efficient the effect of the limited enantioselective inhibition reaction (III) is for the transition of a racemic final solution into a chiral one. Moreover it can also lead to SMSB in the case of a nonexergonic transformation for reactions (I) and (II).

SMSB can also occur if the enantioselective autocatalysis (II) is slower than the uncatalyzed reaction (I) but the transformation is highly exergonic. An example of this is presented in Figure 3. In this case, the kinetics of the transformations (I) and (II) does not exhibit the convex time evolution of the final products, the signature of autocatalysis, but rather a concave curve characteristic of a nonautocatalytic process (linear time-scale).^[17] However, the deracemization stage shows the autocatalytic signature of a convex time evolution of L and D. The chemical significance of this is that SMSB does not require a high selectivity of reaction (II) compared to (I). Notice that the underlying problem in the topic of absolute asymmetric synthesis is the small number of known autocatalytic reactions. The results in the example of Figure 3 suggest that if an enantioselective autocatalysis (II) is "obscured" by the nonautocatalytic reaction (I), a situation that is not easy to detect by kinetic measurements, the system would still be capable of absolute asymmetric synthesis on coupling to a reaction of type (III).

3.2.1. Estimation of the Reaction Parameters Leading to Regions Undergoing SMSB

From a chemical point of view the search for chiral regions in parameter space needs to be made taking into account the chemical reaction parameters, that is, ten that specify the system and not the five dimensionless parameters used for the stability analysis of the racemic and chiral solutions. However, in the search for regions leading to SMSB, a helpful approximation is to first use reaction parameters leading to $g < 1$ and to an effective equilibrium constant at reaction (III) smaller than the equilibrium constant at reaction (II), that is [Eq. (14)]:

$$[X]k_{-3} > k_{-2} \text{ and } \frac{[Y]k_3}{[X]k_{-3}} < \frac{k_3}{k_{-2}} \quad (14)$$

This is because $0 < g < 1$ and at large C values and $u \approx 0$, that is, a relatively high exergonic transformation, then the g_{crit} value (18) approximates to [Eq. (15)]:

$$g_{\text{crit}} \rightarrow \frac{1 - \frac{r}{h}}{1 + \frac{3r}{h}} = \frac{g - \frac{k_3}{k_2}}{g + \frac{3k_3}{k_2}} \quad (15)$$

A second step, when u is not zero and $C < C_{\text{min}}$, is to increase the exergonicity of reactions (I) and (II), which leads to the decrease of u and to an increase in the term $[X]k_{-3}$, which has the effect not only to decrease g but also to increase the dimensionless C value, that is, allowing $C > C_{\text{min}}$ [Eq. (11)]. However, in some cases the change from the chiral to racemic final state is paradoxical. For example, at the conditions of the chiral outcome of Figure 2 in the case of nonexergonic transformations, the increase of efficiency of the enantioselective autocatalysis, upon changing k_2 and k_{-2} to 1×10^2 , leads to the racemic outcome. This is a consequence of the cyclic character of the network studied here, in which an overflow of reaction (II) takes the system out of the metastable region leading to the SMSB.

3.2.2. Conversion of the Cyclic LES Model into a Frank Model

If reaction (III) is replaced by the reaction [Eq. (16)]:

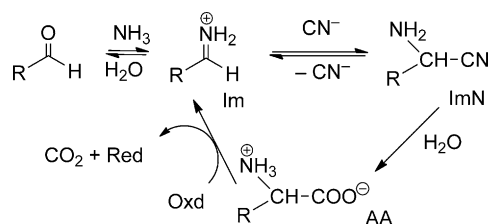


the cyclic LES system transforms into a cyclic Frank-like reaction network, which can also be solved exactly for the asymptotic analytic solutions. Details of this and of the ability of LES and Frank models for SMSB in cyclic reaction networks will be discussed elsewhere. Herein, we only point out that for all simulated examples (Figures 2 and 3) for the same reaction parameters but using the above mutual inhibition reaction [Eq. (16)] instead of reaction (III), no significant differences were observed between both models. All cases exhibit the transition between the chiral and racemic states for the same reaction parameters and lead to similar final concentrations. This suggests that the importance of the inverse reaction in (III) and of the (forward) mutual inhibition (III_{hetero}) in the Frank model is due to the decrease of racemic composition regardless of whether or not it leads to the partial or total destruction of a racemic mixture.

4. Discussion

4.1. Speculative Scenario for a Cyclic Network

To bring these results closer to applied chemistry topics and to the emergence of chirality in prebiotic chemistry, in what follows we discuss a speculative but plausible scenario. We assume as possible a cyclic reaction network (Scheme 2) composed of the Strecker amino acid synthesis and the Strecker amino acid degradation. The cycle is composed of: 1) CN^- addition to an imino (Im) species, a process that, according to the experimental evidence, is nonexergonic or very slightly exergonic; 2) exergonic hydrolysis of the α -amino nitrile (ImN) to the α -amino acid (AA); and 3) its exergonic oxidative decarboxylation towards the initial imino derivative. The feasibility of



Scheme 2. Plausible cyclic reaction network composed of the Strecker α -amino acid synthesis and Strecker amino acid degradation.

such a system depends on the simultaneous presence of these three different reactions. We assume that the CN^- addition and the imino nitrile hydrolysis may occur simultaneously if a specific heterocatalyst acts on one of the two reactions. With respect to the Strecker degradation, it can be promoted by many oxidants.^[18] Some of these oxidants, for example, α -dicarbonyl compounds, can act selectively on the oxidative degradation of the amino acid without oxidizing the imine or the carbaldehyde groups.

The system simulated here by numerical integrations is simplified by the suppression of the imine formation from the carbaldehyde and by assuming constant concentrations of the nucleophile (CN^-) and of the redox pair of the Strecker degradation.

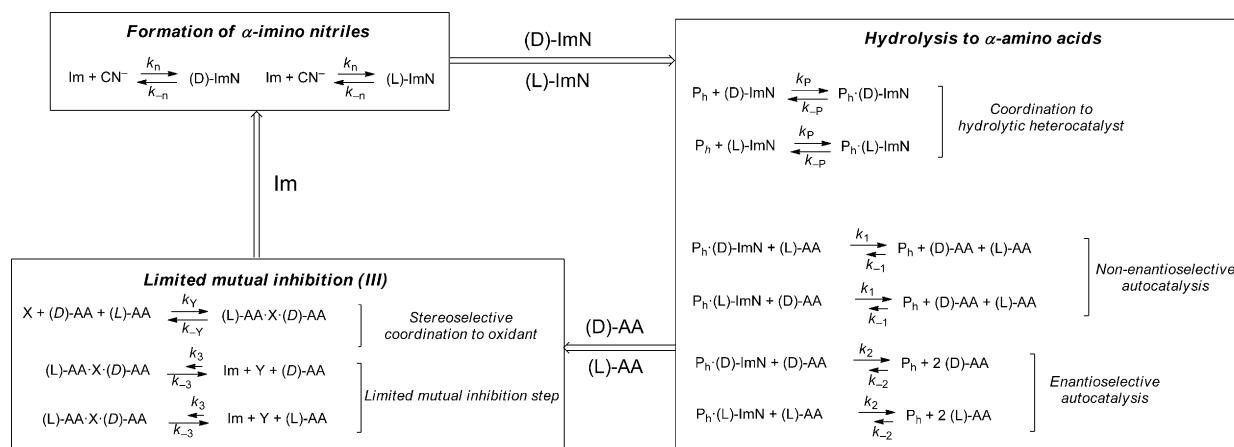
The necessary enantioselective autocatalysis could occur either at the imino nitrile formation or at its hydrolysis to the amino acid. Furthermore, the oxidative fragmentation must also be stereospecific upon a racemic mixture of the amino acid. Hence, a reasonable scenario for SMSB should imply the presence of catalytic coadjuvants leading to stereospecific transformations.

4.1.1. Stereoselectivity Requirements for a Reasonable Applied Chemistry Scenario

The Strecker degradation of Scheme 3 can play the role of the limited inhibition step (III) of a LES network, which implies the stereospecific decomposition of a racemic mixture of the

amino acid enantiomers. This could be achieved by an oxidant (X) able to act after the heterochiral coordination with both enantiomers. In spite of the fact that the oxidative amino acid decarboxylation, in the absence of an enzyme, is a reaction leading to many byproducts, here we assume that in a prebiotic scenario an inorganic solid oxidant (e.g. chiral but in racemic composition) may assume such a catalytic role. In particular, we hypothesize that in the presence of a suitable catalyst, a racemic mixture of amino acids undergoes a specific oxidation so that only one enantiomer is transformed into the achiral starting imine that can reenter the autocatalytic cycle. Clearly, this does not exclude the presence of the direct oxidation, but this must be much slower than the stereospecific oxidation.

Enantiospecific autocatalysis is a rare reaction. Hence, we assume that it may occur after the coordination of the enantiomers with a stereospecific "catalyst". The role of such a catalytic coadjuvant may be played by a clay material, such as a phyllosilicate. Phyllosilicates are sheet silicate materials that can act as catalysts in organic reactions and have been proposed as catalysts in abiotic chemistry scenarios.^[19–21] Furthermore, stereoselective effects on the separation of homochiral and racemic mixtures have been reported^[21,22] and some experimental reports claim enantioselective effects. In fact, some phyllosilicates, for example kaolinite, are intrinsically chiral; therefore, significant enantioselective effects could be expected. Notice that this agrees with several reports on the stereoselective effects of amino acid layered architectures in the origin of biological homochirality.^[23] Therefore, we assume that a heterogeneous catalyst supports the enantioselective autocatalytic reaction. In one of the examples (Scheme 3 and Figure 4) we simply assume that the heterocatalyst is present in constant concentration and yields enantioselective autocatalysis together with a less important nonenantioselective autocatalysis. In another example (Supporting Information) we assume for the catalytic material an initial concentration of achiral sites leading to the nonenantioselective autocatalysis, and a racemic mixture of chiral sites leading to the enantioselective autocatalysis. In summary, enantio- and stereoselective effects can be



Scheme 3. Kinetic scheme of LSE network of Strecker amino acid synthesis coupled to Strecker degradation with the enantioselective autocatalysis at the stage of the imino nitrile (ImN) hydrolysis to amino acid (AA) and stereospecific oxidative degradation. Both stereoeffects occur through the coordination to heterocatalyst coadjuvants. P_n = phyllosilicate heterocatalyst.

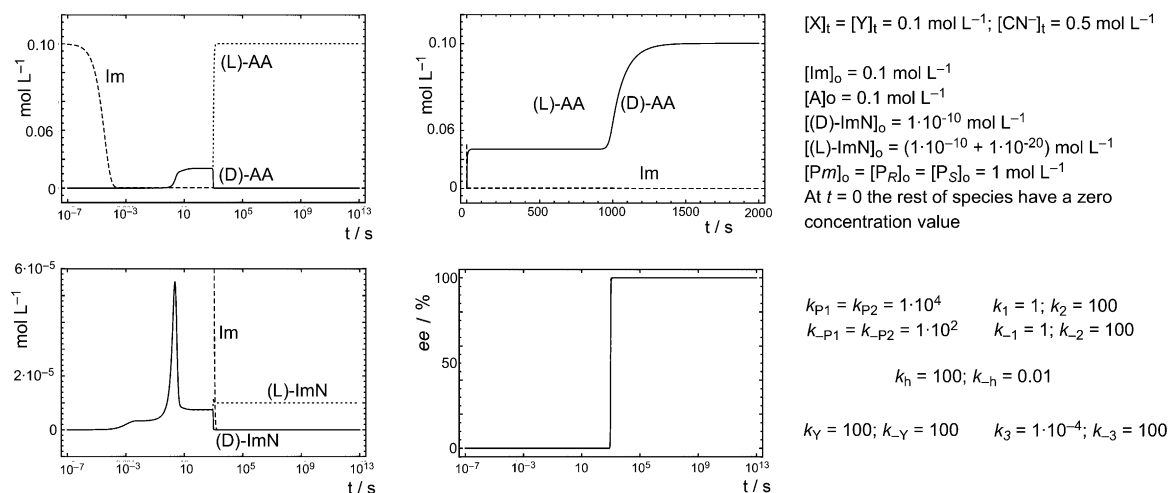


Figure 4. Numerical simulation of the reaction network of Scheme 3 for constant concentrations of the nucleophile CN^- , the redox pair X/Y, and the heterocatalyst P_n . The final ee values for ImN and AA show different chiral signs but very different ee values that, together with the very different final concentrations of ImN and AA, yield a total ee not different, between the limits of any analytical observation, from that of AA.

assumed to be able to occur in a realistic scenario thanks to the presence of racemic mixtures of chiral cocatalysts.

The autocatalytic reactions can occur either in the nucleophilic addition of cyanide ion to the imino compound or in the hydrolysis to the amino acid. In the example of Scheme 3 and Figure 4, the autocatalysis occurs at the hydrolysis stage and in the example of the Supporting Information, in the imino nitrile formation. The first network is a more plausible model for a prebiotic scenario^[24] taking into account the role of water in the catalytic properties of phyllosilicates in hydrolytic reactions.^[20d]

4.2. Examples of SMSB

The SMSB takes place if the initial ee, used to simulate a statistical chiral fluctuation, appears at either the imino nitrile or the amino acid stage. When this ee occurs simultaneously at these two stages, the sign of the final ee is that of the initial ee for the stage at which the autocatalysis occurs, that is, the amino acid hydrolysis in the example of Figure 4 and the imino nitrile formation in the example of the Supporting Information.

In the case of autocatalysis in the formation of imino nitrile (Supporting Information), the final chiral sign is the same for imino nitrile and amino acid and the corresponding ee values are similar. In the case of autocatalysis in the amino acid hydrolysis (Scheme 3 and Figure 4), the final ee in the imino nitrile composition is very low ($\approx 0.001\%$ in the example of Figure 4) and of opposite sign to that of the amino acid ee ($\approx 100\%$). This fact, together with the very low concentration of the imino nitrile, does not lead to a detectable difference of the total ee of the mixture if it is compared with that of the amino acid alone (see Figure 4).

In the case that the amino acid degradation is converted to a Frank-like mutual inhibition [Eq. (16)], for the same reaction parameters the final ee values and species concentrations are similar to those of the LES model. Therefore, the results point to the ability for SMSB in an autocatalytic cycle either in a LES

or in a Frank reaction network, in a cycle containing reaction (III) of Scheme 1 or that of Equation (16).

If SMSB is able to amplify the ee fluctuation values below the expected statistical deviations from the ideal racemic composition, the very small stability differences between D and L enantiomers, which result from the effect of the weak force [energy differences due to the violation of parity (PVED)]^[25] are also able to take the system out from the metastable racemic state to a chiral state of definite sign. Figure 5 shows a simulation of this using the same reaction parameters as those of Figure 4 but an initial ideal racemic composition. The simulation takes into account changes in the rate constants because of the higher stability of the L enantiomers than the D enantiomers (PVED $\approx 1 \text{ fJ mol}^{-1}$, that is, $\Delta k \approx 1 \times 10^{-20}$). Although any change of PVED at the ImN or AA allows the bifurcation towards the chiral state, the important stage that determines the final sign is now reaction (III), for which the higher stability of the L enantiomer determines that the final ee is that of the L enantiomer.

However, notice that the study of the cooperative and/or competitive interactions between chiral statistical fluctuations,

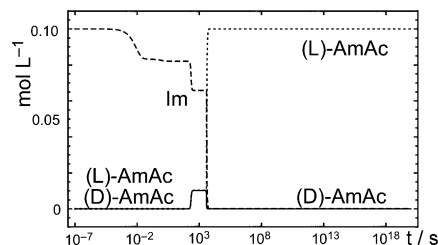


Figure 5. Example of SMSB on the system of Scheme 3 due to PVED ($\approx -1 \text{ fJ mol}^{-1}$ for the L enantiomers). The diagram corresponds to a numerical simulation using the same reaction parameters as those of Figure 4, but an initial racemic composition. Different rate constants for D and L enantiomers were assumed for k_{-n} , k_p , k_{-p} , k_3 , k_{-2} , and k_3 , that is, the corresponding values for the D enantiomers were those of Figure 4 and for the L enantiomers the same but modified by the value -1×10^{-20} .

at different stages of the coupled reactions, and the rate constant differences induced by PVED requires a quite different methodology from the assumptions implicit in chemical kinetics (see, for example, ref. [26]).

5. Conclusions

The results presented herein show, in agreement with a previous report on this topic,^[27] that absolute asymmetric synthesis based on an enantioselective autocatalytic reaction can occur both in conditions disabling the achievement of thermodynamic LES and in Frank models in a cycle of coupled reactions and in equilibrium. In our results such a cycle is constituted by an enantioselective autocatalysis and a reaction implying the decrease of chiral compounds through a heterochiral stereospecific interaction between enantiomers. A chiral stationary state, as the more stable state of the system, may be obtained if the reaction flow is driven by the input of an external reactant (the redox pair X/Y in the examples discussed).

Some analogies of such a cycle with those of the proposals of primordial autocatalytic cycles^[28] suggest that the quest for the emergence of chirality in chemical evolution should be jointly studied with that of these premetabolic cycles.

Acknowledgements

C.B. acknowledges a Calvo-Rodés predoctoral contract from INTA.

Keywords: asymmetric amplification · asymmetric synthesis · autocatalysis · chirality · enantioselectivity

- [1] a) K. Soai, T. Shibata, H. Morioka, K. Choji, *Nature* **1995**, *378*, 767–768; b) K. Soai, T. Kawasaki, *Top. Curr. Chem.* **2008**, *284*, 1–33; c) T. Kawasaki, Y. Matsumura, T. Tsutsumi, K. Suzuki, M. Ito, K. Soai, *Science* **2009**, *324*, 492–495.
- [2] a) C. Viedma, *Phys. Rev. Lett.* **2005**, *94*, 065504; b) W. L. Noorduin, T. Izumi, A. Millemaggi, M. Leeman, H. Meekes, W. J. P. van Enckevort, R. M. Kellogg, B. Kaptein, E. Vlieg, D. G. Blackmond, *J. Am. Chem. Soc.* **2008**, *130*, 1158–1159; c) W. L. Noorduin, E. Vlieg, R. M. Kellogg, B. Kaptein, *Angew. Chem.* **2009**, *121*, 9778–9784; *Angew. Chem. Int. Ed.* **2009**, *48*, 9600–9606; d) C. Viedma, W. L. Noorduin, J. E. Ortiz, T. de Torres, P. Cintas, *Chem. Commun.* **2011**, *47*, 671–673; e) Z. El-Hachemi, J. Crusats, J. M. Ribó, J. M. McBride, S. Veintemillas-Verdaguer, *Angew. Chem.* **2011**, *123*, 2407–2411; *Angew. Chem. Int. Ed.* **2011**, *50*, 2359–2363; f) C. Viedma, P. Cintas, *Chem. Commun.* **2011**, *47*, 12786–12788.
- [3] An example of this is the Tsogoeva report^[4] on SMSB in Mannich and aldolic reactions. The interest of the report lies in the questions that arise on obtaining temporary or permanent chiral outcomes independently of whether the yields are too low to be of interest for organic synthesis.
- [4] M. Mauksch, S. B. Tsogoeva, S. Wei, I. M. Martynova, *Chirality* **2007**, *19*, 816–825.
- [5] In fact, the addition of an achiral single molecule to the faces of a prochiral C=O or C=N group is a SMSB process, but due to the role of the very high numbers, this can only lead to a racemic mixture. However, in chemical networks, such as those discussed here, an initially achiral macroscopic system can achieve a final chiral state either by kinetic control or by thermodynamic control in the case of a final stationary state.
- [6] F. C. Frank, *Biochim. Biophys. Acta* **1953**, *11*, 459–463.
- [7] a) J. R. Islas, D. Lavabre, J.-M. Grevy, R. H. Lamoneda, H. R. Cabrera, J.-C. Micheau, T. Buhse, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13743–13748; b) D. Lavabre, J.-C. Micheau, J. R. Islas, T. Buhse, *J. Phys. Chem. A* **2007**, *111*, 281–286; c) D. Lavabre, J.-C. Micheau, J. R. Islas, T. Buhse, *Top. Curr. Chem.* **2008**, *284*, 67–96; d) J. Crusats, D. Hochberg, A. Moyano, J. M. Ribó, *ChemPhysChem* **2009**, *10*, 2123–2131; e) C. Blanco, M. Stich, D. Hochberg, *Chem. Phys. Lett.* **2011**, *505*, 140–147.
- [8] a) V. Avetisov, V. Goldanskii, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11435–11442; b) K. Iwamoto, *Phys. Chem. Chem. Phys.* **2002**, *4*, 3975–3979; c) J. C. Micheau, C. Coudret, J.-M. Cruz, T. Buhse, *Phys. Chem. Chem. Phys.* **2012**, *14*, 13239–13248.
- [9] In the Frank model this is due to the “mutual inhibition” between enantiomers (III_{hetero} in Scheme 1, left-hand side), while in LES it is caused by the reverse reaction of the nonenantioselective autocatalysis (III_{rac} and III in Scheme 1, right-hand side).
- [10] a) W. H. Mills, *Chem. Ind.* **1932**, *51*, 750–759; b) V. I. Goldanskii, V. V. Kuzmin, *Z. Phys. Chem.* **1988**, *269*, 216–274; c) K. Mislow, *Collect. Czech. Chem. Commun.* **2003**, *68*, 849–864.
- [11] J. Jacques, A. Collet, S. H. Wilen, *Enantiomers, Racemates and Resolutions*, Wiley, New York, **1981**, p. 81.
- [12] D. G. Blackmond, O. K. Matar, *J. Phys. Chem. B* **2008**, *112*, 5098–5104.
- [13] J. M. Ribó, D. Hochberg, *Phys. Lett. A* **2008**, *373*, 111–122.
- [14] As “resistant chiral amplification” we understand a reaction dynamics that achieves an ee plateau that maintains its ee value, in the limits of the experimental detection, and when it eventually racemizes, it does so abruptly.
- [15] a) J. M. Ribó, J. Crusats, A. El-Hachemi, A. Moyano, C. Blanco, D. Hochberg, *Astrobiology* **2013**, *13*, 132–142; b) C. Blanco, J. M. Ribó, J. Crusats, Z. El-Hachemi, A. Moyano, D. Hochberg, *Phys. Chem. Chem. Phys.* **2013**, *15*, 1546–1556.
- [16] In the examples of Figure 2 the effect is achieved by changing only k_{-3} , but it can also be obtained by changing the [S] value.
- [17] R. Plasson, A. Brandenburg, L. Jullien, H. Bersini, *J. Phys. Chem. A* **2011**, *115*, 8073–8085.
- [18] A. Schönberg, R. Moubacher, *Chem. Rev.* **1952**, *50*, 261–277.
- [19] The possible role of mineral surfaces in the selection of organic compounds and as catalysts in prebiotic processes was already presented in the first proposals of chemical evolution as the physical basis of life, and has been the subject of elaborate mineral-based scenarios on the origin of life.^[20]
- [20] a) A. G. Cairns-Smith, H. Hartman, *Clay Minerals and the Origin of Life*, Cambridge University Press, Cambridge, **1986**; b) “Clay Minerals and the Origin of Life in Developments in Clay Science”: A. Brack, *Handbook of Clay Science, Vol. 1* (Eds.: F. Bergaya, K. G. Theng, G. Lagaly), Elsevier, Oxford, **1986**; c) J. Bujdak, H. Slosiarikova, N. Textler, M. Schwendiger, B. M. Rode, *Monatsh. Chem.* **1994**, *125*, 1033–1039; d) *Pillared Clays and Related Catalysts* (Eds.: A. Gil, S. A. Korilli, R. Trujillano, M. A. Vicente), Springer, Berlin, **2010**.
- [21] L. Guijarro, M. Yus, *The Origin of Chirality in the Molecules of Life*, RSC, Cambridge, **2009**, pp. 129–132.
- [22] a) S. C. Bondy, M. Harrington, *Science* **1979**, *203*, 1243–1244; b) A. Yamagishi, H. Sato, *Clays Clay Miner.* **2012**, *60*, 411–419.
- [23] I. Weissbuch, M. Lahav, *Chem. Rev.* **2011**, *111*, 3236–3267.
- [24] The relationship between the presence of water and the formation of phyllosilicates (serpentinization) is an accepted point in the evolution of the silica materials in planetesimal bodies.
- [25] a) A. J. MacDermott, *Origins Life Evol. Biosphere* **1995**, *25*, 191–199; b) W. A. Bonner, *Origins Life Evol. Biosphere* **1999**, *29*, 615–623; c) M. Gottselig, M. Quack, *J. Chem. Phys.* **2005**, *123*, 084305.
- [26] a) D. K. Kondepudi, I. Prigogine, G. Nelson, *Phys. Lett. A* **1985**, *111*, 29–32; b) V. A. Avetisov, V. V. Kuzmin, S. A. Anikin, *Chem. Phys.* **1987**, *112*, 179–187; c) G. Lente, *J. Phys. Chem. A* **2004**, *108*, 9475–9478; d) J. M. Cruz, P. Parmananda, T. Buhse, *J. Phys. Chem. A* **2008**, *112*, 1673–1676.
- [27] R. Plasson, H. Bersini, A. Commeyras, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16733–16738.
- [28] a) G. Wächtershäuser, *Microbiol. Rev.* **1988**, *52*, 452–484; b) G. Wächtershäuser, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 200–204; c) H. J. Morowitz, J. D. Kostelnik, J. Yang, G. D. Cody, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7704–7708; d) A. Eschenmoser, *Tetrahedron* **2007**, *63*, 12821–12844.

Received: April 8, 2013

Published online on July 2, 2013

Mirror symmetry breaking with limited enantioselective autocatalysis and temperature gradients: a stability survey

Cite this: *Phys. Chem. Chem. Phys.*, 2013, **15**, 1546

Celia Blanco,^a Josep M. Ribó,^{bc} Joaquim Crusats,^{bc} Zoubir El-Hachemi,^{bc} Albert Moyano^b and David Hochberg^{*a}

We analyze limited enantioselective (LES) autocatalysis in a temperature gradient and with internal flow/recycling of hot and cold material. Microreversibility forbids broken mirror symmetry for LES in the presence of a temperature gradient alone. This symmetry can be broken however when the autocatalysis and limited enantioselective catalysis are each localized within the regions of low and high temperature, respectively. This scheme has been recently proposed as a plausible model for spontaneous emergence of chirality in abyssal hydrothermal vents. Regions in chemical parameter space are mapped out in which the racemic state is unstable and bifurcates to chiral solutions.

Received 3rd October 2012,
Accepted 23rd November 2012

DOI: 10.1039/c2cp43488a

www.rsc.org/pccp

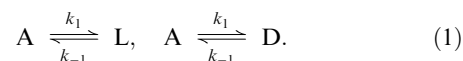
1 Introduction

Recent experimental reports on the deracemization of racemic mixtures of crystals and on the crystallization from boiling solutions^{1–6} are striking examples of novel scenarios for spontaneous mirror symmetry breaking (SMSB) of compounds for which the homochiral interactions are favored over the heterochiral ones. In other words, these are reactions that cannot be explained by Frank-like mechanisms, in which the heterochiral interaction is the favored one. Despite some controversy about the actual mechanisms responsible for the SMSB in these situations, the experimental reports all coincide in that the final state is stationary: mechano-stationary in the case of wet grinding of racemic mixtures of crystals^{3–5} and the presence of temperature gradients in the case of deracemization and crystallization in boiling solutions.^{1,2} The above reports stress the fact that the racemic conglomerate crystal mixtures are deracemized under experimental conditions where *chemical equilibrium is not possible, i.e.*, they require specific energy input to only some of the species of the system (as in the

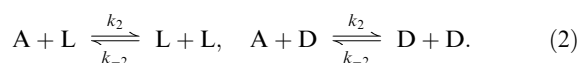
crystal grinding experiments) or else a non-uniform temperature distribution.

The point of departure of the present paper is that if limited enantioselectivity⁷ under experimental conditions of closed systems with a uniform distribution of temperature and energy inexorably yields a final racemic state,⁸ then can SMSB occur in scenarios of non-uniform temperature distributions? As it turns out, this is in fact possible in the case that, in addition to a non-uniform temperature distribution, the systems possess a compartmentalization of the enantioselective and the non-enantioselective autocatalyses.

The emergence of chirality in enantioselective autocatalysis of compounds which do not follow Frank-like schemes is investigated here for the limited enantioselectivity (LES) model composed of coupled enantioselective and non-enantioselective autocatalyses. The basic model⁷ is defined by the following chemical transformations. Production of chiral compounds L,D from an achiral substrate A:



Autocatalytic production:



Limited enantioselectivity:



^a Centro de Astrobiología (CSIC-INTA), Ctra. Ajalvir Km. 4, 28850 Torrejón de Ardoz, Madrid, Spain. E-mail: blancodtc@cab.inta-csic.es, hochbergd@cab.inta-csic.es

^b Department of Organic Chemistry, University of Barcelona, c. Martí i Franquès 1, 08028 Barcelona, Catalonia, Spain. E-mail: jmribo@ub.es, j.crusats@ub.es, zelhachemi@ub.es, amoyano@ub.es

^c Institute of Cosmos Science, University of Barcelona (IEEC-UB), c. Martí i Franquès 1, 08028 Barcelona, Catalonia, Spain

In contrast to the Frank model, LES is able to account for two important facts: namely, (i) the enantioselectivity of any chiral catalyst is limited because of the third reaction, eqn (3), and (ii) the kinetic link between mirror conjugate processes arises from the reversibility of the catalytic stage.⁷ The inverse reaction of the non-enantioselective autocatalysis (reaction (3)) substitutes for the mutual inhibition reaction in the Frank model or formation of the heterodimer ($L + D \rightarrow P$). Earlier reports had claimed spontaneous mirror symmetry breaking (SMSB) in LES, but this cannot occur in either open or closed systems with a *uniform* temperature distribution. The obstacle comes from microreversibility, where $K(T)$ is the temperature dependent equilibrium constant:

$$\frac{k_i}{k_{-i}} = K(T), \quad (1 \leq i \leq 3). \quad (4)$$

The condition for the instability of the racemic state is that

$$0 < g < \frac{1-w}{1+3w} < 1, \quad (5)$$

where $g = \frac{k_{-2}}{k_{-3}}$ and $w = \frac{k_3}{k_2}$.⁸ From eqn (5), we must have $1 - w > 0$

so that $1 > w$. But from eqn (4), $\frac{k_2}{k_3} = \frac{k_{-2}}{k_{-3}}$, which is incompatible with $1 > w$ and $g < 1$. This is the situation for open systems. For closed systems, it can be shown⁸ that the racemic state is unstable provided that

$$0 < g < g_{\text{crit}}^{\text{closed}} < g_{\text{crit}}^{\text{open}} = \frac{1-w}{1+3w} < 1, \quad (6)$$

where the critical parameter in a closed system is always bounded above by the corresponding one for open systems, and approaches the latter from below in the limit of large total concentrations C , that is: $g_{\text{crit}}^{\text{closed}} \rightarrow g_{\text{crit}}^{\text{open}}$. But eqn (6) is also incompatible with eqn (4). So the racemic state is always asymptotically stable in this scheme for both open and closed systems held at a uniform temperature. Therefore there can be no asymptotically stable chiral outcome in this model.

Nor can the LES model lead to SMSB in closed systems even with a stationary non-uniform temperature distribution. On the other hand, numerical simulations of chemical kinetics in a two-compartment model⁹ demonstrate that SMSB may occur if both catalytic reactions (2) and (3) are spatially separated at different temperatures in different compartments but coupled under the action of a continuous internal flow. Under such conditions the system can evolve, for certain reaction and system parameters, towards a chiral stationary state, *i.e.*, the system is able to reach a bifurcation point leading to SMSB.

This is an appealing result since numerical simulations using reasonable chemical parameters suggest that an adequate scenario for such an LES-based SMSB would be that of abyssal hydrothermal vents, by virtue of the typical temperature gradients found there and inorganic solids mediating chemical reactions in an enzyme-like role. We therefore proposed⁹ that a natural prebiotic scenario for such emergence of chirality is

that of abyssal hydrothermal vents and volcanic plumes^{10–13} which do have the adequate temperature gradients and contain solids, as for example clays, which have been proposed by several authors^{14–18} as catalysts in the prebiotic synthesis of organic compounds.

In view of the above, this paper deals with an analytic/numerical study of the conditions leading to the instability of the ideal racemic composition for the LES model with compartmentalized catalyses (2) and (3) in regions held at different temperatures. The two-compartment model (Section 2) is already sufficiently involved as to make deriving general analytic stability results a near impossible task. We thus focus our efforts on analyzing properties of the racemic fixed point; the analytic conditions for its linear stability can be set up and then tested in numerical domains. The direct study of the stationary chiral solutions is substantially more complicated and algebraically unwieldy. Nevertheless, the characterization of a racemic state as unstable necessarily implies that the system evolves to a state of non-racemic composition. Such a nonracemic state could be stable, chaotic or oscillatory. However, the many numerical SMSB tests performed for reasonable reaction parameters and under the thermodynamic constraints imposed by the principle of microreversibility and the temperature dependencies of the reaction rate constants (see Section 3) have in all cases led to a stable chiral state. These stable states are scalemic mixtures of enantiomeric excesses (ee) whose values depend on the parameters of the phase diagram. We first review the impossibility of SMSB in LES with a stable temperature gradient. We then prove that SMSB in such a system is possible when the two catalyses (2) and (3) are compartmentalized (localized) in different temperature regions connected by an internal flow of material.

2 LES with a temperature gradient

Consider the LES scheme in a temperature gradient: this permits the reaction rates to vary spatially in the system, from one place to another, and might provide a way to achieve mirror symmetry breaking. The reverse reaction of (3) in one region (compartment) could be faster than the reverse of (2) in the other. And then mixing could bring the hot and cold material into contact. We model this by a closed two-compartment system with volumes V and V^* ; each compartment held at a uniform temperature, $T^* > T$, and internally coupled by a constant internal flow or recycling. The concentrations and reaction rates for the second compartment are labeled by an asterisk, see Fig. 1. We extend the basic LES scheme, eqn (1)–(3), by including the terms corresponding to the internal flow. The internal “flow” parameter a has units of volume per time; V and V^* denote the volumes of each compartment. The corresponding rate equations for the two-compartment system have been derived in ref. 9.

2.1 New variables

For setting up a stability analysis it is convenient to employ the sums and differences of the concentrations $\chi = [L] + [D]$, $y = [L] - [D]$,

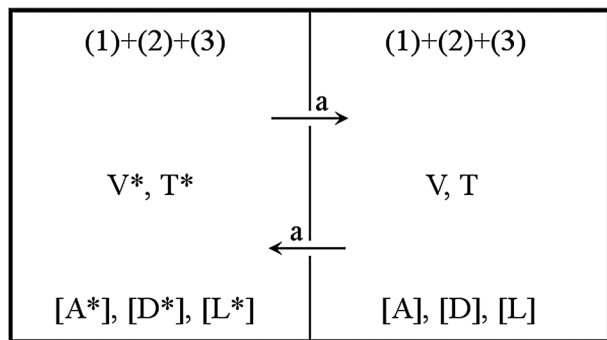


Fig. 1 Limited enantioselectivity (LES) in two compartments with volumes V and V^* each held at different temperatures $T^* > T$ and interconnected by an internal flow a of the material. See ref. 9.

$\chi^* = [L^*] + [D^*]$ and $y^* = [L^*] - [D^*]$. In terms of these, the rate equations take the following form:

$$\frac{d[A]}{dt} = -2k_1[A] - (k_2[A] + k_3[A] - k_{-1})\chi + \frac{k_{-2}}{2}(y^2 + \chi^2) + \frac{k_{-3}}{2}(\chi^2 - y^2) + \frac{a}{V}([A^*] - [A]), \quad (7)$$

$$\frac{d\chi}{dt} = +2k_1[A] + (k_2[A] - k_{-1})\chi + k_3[A]\chi - k_{-2}\frac{(\chi^2 + y^2)}{2} - k_{-3}\frac{(\chi^2 - y^2)}{2} + \frac{a}{V}(\chi^* - \chi), \quad (8)$$

$$\frac{dy}{dt} = +(k_2[A] - k_{-1})y - k_{-2}\chi y - k_3[A]y + \frac{a}{V}(y^* - y), \quad (9)$$

$$\frac{d[A^*]}{dt} = -2k_1^*[A^*] - (k_2^*[A^*] + k_3^*[A^*] - k_{-1}^*)\chi^* + k_{-2}^*(y^{*2} + \chi^{*2}) + \frac{k_{-3}^*}{2}(\chi^{*2} - y^{*2}) + \frac{a}{V^*}([A] - [A^*]), \quad (10)$$

$$\frac{d\chi^*}{dt} = +2k_1^*[A^*] + (k_2^*[A^*] - k_{-1}^*)\chi^* + k_3^*[A^*]\chi^* - k_{-2}^*\frac{(\chi^{*2} + y^{*2})}{2} - k_{-3}^*\frac{(\chi^{*2} - y^{*2})}{2} + \frac{a}{V^*}(\chi - \chi^*), \quad (11)$$

$$\frac{dy^*}{dt} = +(k_2^*[A^*] - k_{-1}^*)y^* - k_{-2}^*\chi^*y^* - k_3^*[A^*]y^* + \frac{a}{V^*}(y - y^*). \quad (12)$$

They satisfy the constant mass constraint:

$$V([A] + \chi) + V^*([A^*] + \chi^*) = C, \quad (13)$$

where C is the total conserved mass in the complete two-compartment system.

3 Constraints from Arrhenius–Eyring

Certain thermodynamic relationships hold among the reaction rates in both compartments. These will be used to prove that SMSB is also impossible for the scheme presented in Section 2 in a background temperature gradient. Following this demonstration, we then introduce the variant of LES that can and does lead to SMSB.

From Arrhenius–Eyring, the forward (and reverse) reaction rates for reaction i at temperatures T^* and $T < T^*$ are¹⁹

$$k_i^* = \left(\frac{k_B T^*}{h}\right) e^{-\frac{\Delta G_i(T^*)}{RT^*}}, \quad k_{-i}^* = \left(\frac{k_B T^*}{h}\right) e^{-\frac{\Delta G_{-i}(T^*)}{RT^*}}, \quad (14)$$

$$k_i = \left(\frac{k_B T}{h}\right) e^{-\frac{\Delta G_i(T)}{RT}}, \quad k_{-i} = \left(\frac{k_B T}{h}\right) e^{-\frac{\Delta G_{-i}(T)}{RT}}, \quad (15)$$

here

$$\Delta G_i(T) = \Delta H_i - T\Delta S_i \quad (16)$$

denotes the difference in free energy between the activated state (transition state) and the reactants, while

$$\Delta G_{-i}(T) = \Delta H_{-i} - T\Delta S_{-i}, \quad (17)$$

is the free energy difference between the activated state (transition state) and the products; H and S denote the enthalpy and entropy, respectively. From the above we can obtain a relationship between the forward reaction rates i in each compartment:

$$k_i^* = \left(\frac{T^*}{T}\right) \exp\left(-\frac{\Delta H_i}{R}\left(\frac{1}{T^*} - \frac{1}{T}\right)\right) k_i. \quad (18)$$

Clearly, once the values of the k_i are chosen for the reference compartment at T , we are not free to *independently* choose the reaction rates k_i^* at the higher temperature T^* .

The fundamental microreversibility condition, eqn (4), together with Arrhenius–Eyring implies

$$\frac{k_i}{k_{-i}} = \frac{e^{-\frac{\Delta G_i}{RT}}}{e^{-\frac{\Delta G_{-i}}RT}} = e^{(\Delta G_{-i} - \Delta G_i)/RT} = K(T), \quad (19)$$

$$\Leftrightarrow \Delta G_{-i} - \Delta G_i \equiv \Delta\Delta G, \quad (20)$$

that is, the difference in the free energy differences $\Delta\Delta G$ must be *independent* of i , that is, independent of the specific i th reaction. This implies that the individual double differences in enthalpy and in entropy must also be independent of reaction i , so we must also have

$$\Delta\Delta H = (\Delta H_{-i} - \Delta H_i), \quad (21)$$

$$\Delta\Delta S = (\Delta S_{-i} - \Delta S_i). \quad (22)$$

This also gives us an expression for calculating K :

$$K(T) = e^{\frac{\Delta\Delta H}{RT}} e^{-\frac{\Delta\Delta S}{R}}. \quad (23)$$

The inverse reaction rates are obtained through the constraint eqn (4) as follows:

$$k_{-i} = \frac{k_i}{K(T)}. \quad (24)$$

We also note that if the constraint eqn (4) is satisfied at one specific temperature T , then it will automatically hold at all others, that is

$$\frac{k_i^*}{k_{-i}^*} = K(T^*) = e^{\frac{\Delta\Delta H}{RT^*}} e^{-\frac{\Delta\Delta S}{R}}, \quad (1 \leq i \leq 3). \quad (25)$$

The ratio of the equilibrium constants is given by

$$\frac{K(T^*)}{K(T)} = \exp\left(\frac{\Delta\Delta H}{R}\left(\frac{1}{T^*} - \frac{1}{T}\right)\right). \quad (26)$$

The algebraic intricacy of the model in Section 2 is already such that we are unable to obtain *useful* and manageable analytic closed form expressions for the conditions leading to the instability of the racemic solution. The situation is even worse for obtaining analytic information regarding the possible stationary chiral solutions. We appeal instead to chemically inspired conjectures that can be tested numerically for coherence and compatibility with microreversibility.

First, it is clear that in view of the gradient $T < T^*$, the putative condition, which could conceivably lead to symmetry breaking in the limit of small values of a ,

$$\begin{aligned} k_{-2} < k_{-3} \quad \& \quad k_2 > k_3 \quad \text{at } T \\ \text{and } k_{-2}^* < k_{-3}^* \quad \& \quad k_2^* > k_3^* \quad \text{at } T^*, \end{aligned} \quad (27)$$

is incompatible with the constraints in eqn (23) and (25). This condition (27) is inspired by the observation that for $a \rightarrow 0$, the two compartments are practically isolated from each other and can be treated as approximately independent. These are thus the necessary conditions for obtaining an unstable racemic solution in each compartment (see Section 1). But they are incompatible with microreversibility.

Secondly, the analysis in Section 1 suggests that SMSB might occur when the inverse reaction of (3) in one region is *faster* than the inverse reaction of (2) in the other region. Taking microreversibility into account, the only way this might be achieved is, for example, by arranging for

$$\begin{aligned} k_{-3} < k_{-2} \quad \& \quad k_2 > k_3 \quad \text{at } T \\ \text{and } k_{-2}^* < k_{-3}^* \quad \& \quad k_2^* < k_3^* \quad \text{at } T^*. \end{aligned} \quad (28)$$

But this is forbidden by virtue of eqn (26), which is satisfied by the ratio of the equilibrium constants. So a temperature gradient and internal flow are by themselves not enough to produce a bifurcation. Actually, no spatially varying temperature profile is sufficient, as can be seen by partitioning the closed system into a number of sufficiently small regions within which the local temperature is approximately uniform.

$$A = \begin{pmatrix} \left(-2k_1 - [k_2 + k_3]\bar{\chi} - \frac{a}{V} - \frac{a}{V^*} \right) & \left(-(k_2 + k_3)\bar{A} + (k_{-2} + k_{-3})\bar{\chi} + k_{-1} - \frac{a}{V^*} \right) & -\frac{a}{V} \\ (2k_1 + [k_2 + k_3]\bar{\chi}) & \left((k_2 + k_3)\bar{A} - k_{-1} - (k_{-2} + k_{-3})\bar{\chi} - \frac{a}{V} \right) & \frac{a}{V} \\ -\frac{V}{V^*}(2k_1^* + [k_2^* + k_3^*]\bar{\chi}^*) & \left(\frac{a}{V^*} - \frac{V}{V^*}(2k_1^* + [k_2^* + k_3^*]\bar{\chi}^*) \right) & \left([k_2^* + k_3^*]\left(\frac{C}{V^*} - \frac{V}{V^*}(\bar{A} + \bar{\chi}) - \bar{\chi}^*\right) \right) \end{pmatrix}, \quad (31)$$

4 Temperature gradient and immobilized catalysts

Our working hypothesis is that a *necessary but not sufficient* condition for the instability of the racemic solution is $k_{-3}^* > k_{-2}$ and $k_2 > k_3^*$ in the presence of immobilized catalysts that ensure that $k_{\pm 2}^* = 0$ in one region and $k_{\pm 3} = 0$ in the other.⁹ Other conditions (total system concentration C , the flow rate a , compartment volumes V , V^* , etc.) also come into play for determining the overall instability, in a highly nontrivial and nonlinear fashion.

4.1 Linear stability analysis of the stationary racemic fixed point

Eqn (9) and (12) for $\frac{dy}{dt} = \frac{dy^*}{dt} = 0$ are identically satisfied for the stationary solution $\bar{y} = \bar{y}^* = 0$. We therefore carry out a stability analysis of this racemic fixed point and determine whether the racemic solution (racemic in *both* compartments) is asymptotically stable or unstable. This will depend on the internal *flow parameter* a that characterizes the cycling of hot to cold material between the two compartments. Clearly, if we set $a = 0$, we merely recover two isolated copies of LES in independent closed compartments, each one at a constant temperature, and there can be no mirror symmetry breaking in this situation; the considerations of Section 1 apply.

An algebraic advantage of studying the racemic fixed point $\bar{y} = \bar{y}^* = 0$ is that the five independent concentration fluctuations decouple into two sets of three and two, respectively. This situation is reflected in the structure of the Jacobian matrix J which then reduces to a block-diagonal form with a 3×3 sub-block corresponding to the fluctuations $(\delta A, \delta\chi, \delta\chi^*)$ and a 2×2 sub-block corresponding to $(\delta y, \delta y^*)$ thus:

$$J = \begin{pmatrix} A^{3 \times 3} & D^{3 \times 2} \\ C^{2 \times 3} & B^{2 \times 2} \end{pmatrix} \Rightarrow \begin{pmatrix} A^{3 \times 3} & 0 \\ 0 & B^{2 \times 2} \end{pmatrix}. \quad (29)$$

The temporal evolution of the linearized concentration fluctuations about the racemic fixed point $\bar{y} = \bar{y}^* = 0$ of the kinetic equations is given by

$$\frac{d}{dt} \begin{pmatrix} \delta A \\ \delta\chi \\ \delta\chi^* \end{pmatrix} = A \begin{pmatrix} \delta A \\ \delta\chi \\ \delta\chi^* \end{pmatrix}$$

and

$$\frac{d}{dt} \begin{pmatrix} \delta y \\ \delta y^* \end{pmatrix} = B \begin{pmatrix} \delta y \\ \delta y^* \end{pmatrix}, \quad (30)$$

where the 3×3 array A is given by eqn (31) and the 2×2 array B is given by eqn (32), (see Appendix A)

$$B = \begin{pmatrix} \left(-k_{-1} - k_{-2}\bar{\chi} + (k_2 - k_3)\bar{A} - \frac{a}{V} \right) \frac{a}{V} \\ \frac{a}{V^*} \\ \left((k_2^* - k_3^*) \left(\frac{C}{V^*} - \frac{V}{V^*}(\bar{A} + \bar{\chi}) - \bar{\chi}^* - k_{-1}^* - k_{-2}^*\bar{\chi}^* - \frac{a}{V^*} \right) \right) \end{pmatrix}. \quad (32)$$

The Jacobian matrix J (29) must be evaluated on non-negative stationary solutions $\bar{A} \geq 0$, $\bar{\chi} \geq 0$, $\bar{\chi}^* \geq 0$ corresponding to $\bar{y} = \bar{y}^* = 0$. Then, to assess the stability of the solution, the five

eigenvalues λ_i , $i = 1, 2, \dots, 5$ of the Jacobian matrix must be calculated. If any one of these five eigenvalues is positive (or their real part, if complex) then the solution is unstable.

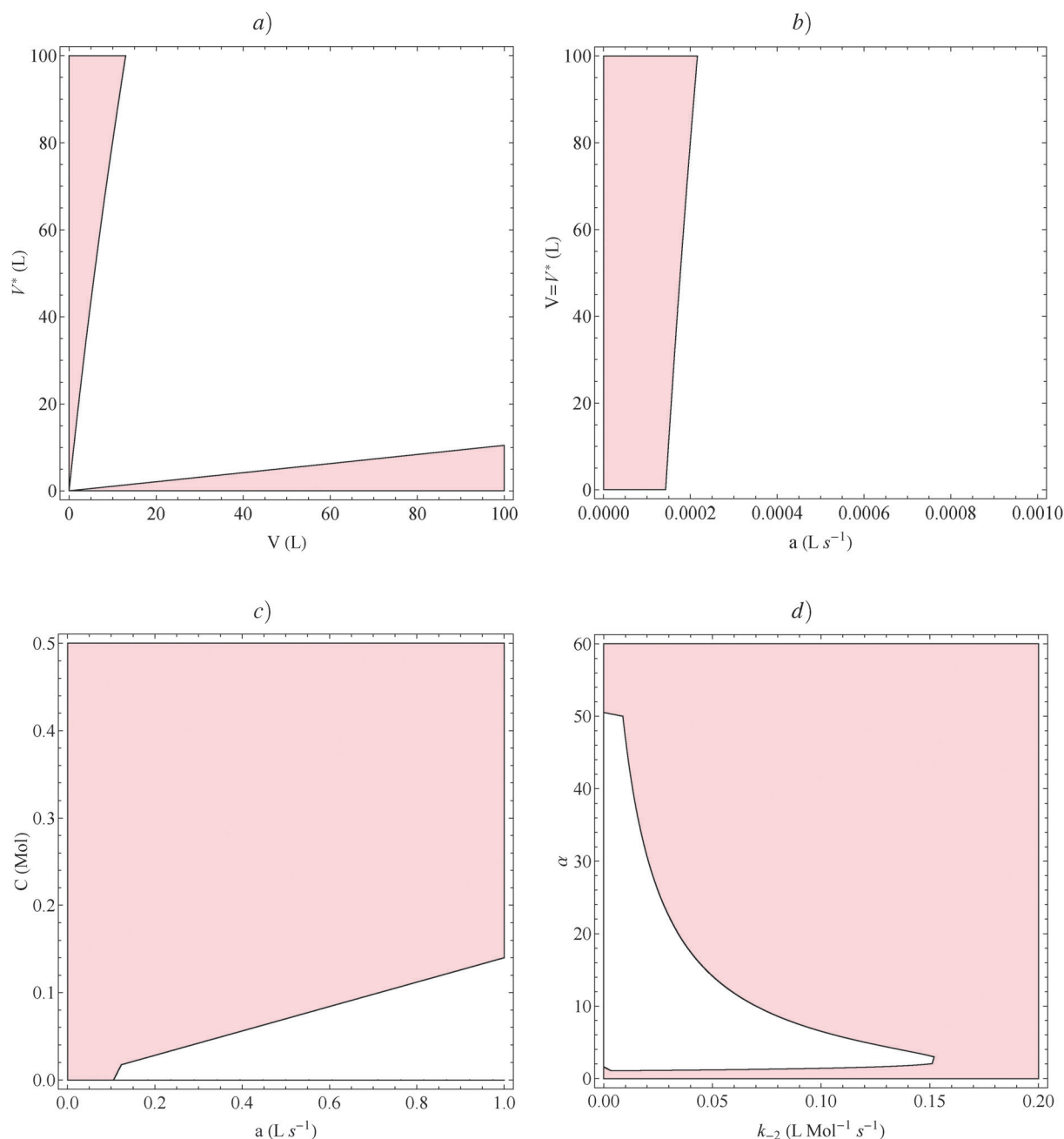


Fig. 2 White zones indicate where the racemic state is *unstable* to perturbations and so bifurcates to a chiral state (color/grayscale, where the racemic state is stable) as a function of the indicated selected pairs of variables: (a) the cold and hot compartment volumes V and V^* , (b) the flow parameter a versus the volume of each compartment, holding $V = V^*$ fixed, (c) the flow parameter a and the total system concentration C , and (d) the reaction rate k_{-2} and α , where $\alpha = k_{-3}^*/k_{-2}$. Except for the specific pair that is varied in (a), (b), (c), and (d), the remainder of values are held fixed at $k_1 = 10^{-9}$, $k_{-1} = 10^{-15}$, $k_1^* = 10^{-7}$, $k_{-1}^* = 10^{-11}$, $k_2 = 10^2$, $k_{-2} = 10^{-4}$, $k_2^* = 0$, $k_{-2}^* = 0$, $k_3 = 0$, $k_{-3} = 0$, $k_3^* = 10^1$, $k_{-3}^* = 10^{-3}$, the volume and flux parameters $V = 10$, $V^* = 10$, $a = 0.1$. The initial concentrations $[A]_0 = 10^{-6}$, $[A^*]_0 = 10^{-6}$, $[L]_0 = 10^{-11}$, $[D]_0 = 10^{-11}$, $[L^*]_0 = 10^{-11} + 10^{-21}$, $[D^*]_0 = 10^{-11}$.

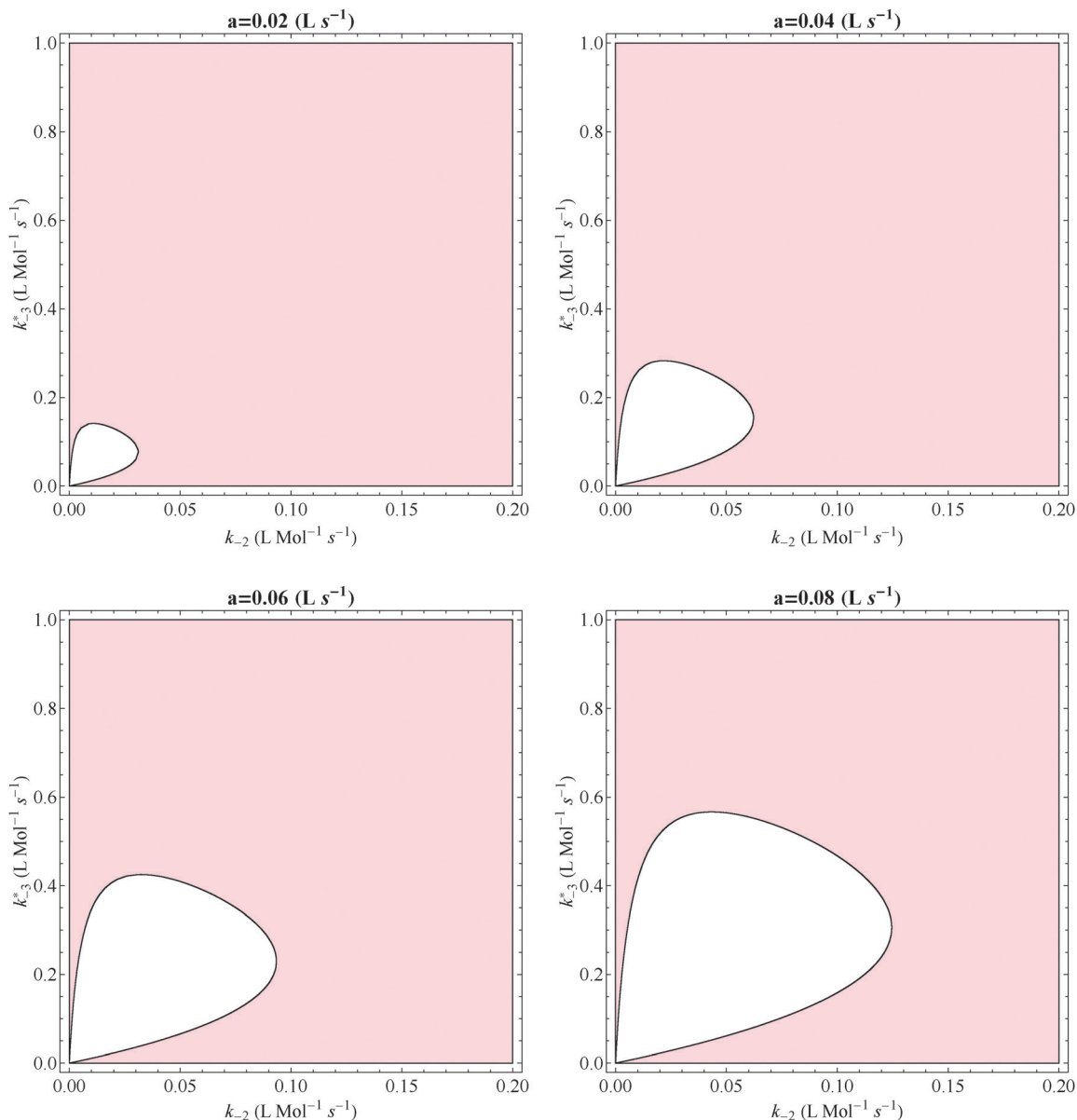


Fig. 3 White zones indicate where the racemic state is *unstable* to perturbations (color/grayscale, where it is stable) for different values of k_{-2} and k_{-3}^* , and for different values of the internal flow parameter a . The area of the region of instability scales up with the flow rate. The remainder of parameters and initial concentrations as in Fig. 2.

This means that the system will evolve to a chiral state, so mirror symmetry will be broken. Only if *all* five of the eigenvalues are negative (real part) we can claim that the solution is stable. Deriving manageable and useful closed form expressions for the eigenvalues λ_i of eqn (31) and (32) is practically impossible, due to the fact that the racemic fixed point solutions \bar{A} , $\bar{\chi}$, $\bar{\chi}^*$ lead to unwieldy expressions (as solutions of coupled quartic equations). On the other hand, direct numerical calculation of the fixed point solutions and their associated eigenvalues is amenable and provides a wealth of information about the dynamic stability of the underlying model, as functions of the chemical rates and the system parameters. We will therefore map out regions of stability/instability in parameter space. We carry out this assuming the immobilized catalyst

from the start, setting k_2^* , k_{-2} to zero in one compartment and k_3 , k_{-3} in the other. Variations in the remaining rate constants are carried out satisfying the constraints in eqn (24) and (25).

We apply a second, independent, stability test which does not require calculation of the eigenvalues: namely, the Routh–Hurwitz (RH) criteria. We derive explicit expressions whose algebraic signs indicate whether the racemic fixed point is stable or unstable. The canonical form of the characteristic polynomial for the complete 5×5 Jacobian matrix, eqn (29), is

$$P(\lambda) = \lambda^5 + a_1\lambda^4 + a_2\lambda^3 + a_3\lambda^2 + a_4\lambda + a_5 = 0. \quad (33)$$

Then (see Appendix B of ref. 20) there are conditions on the coefficients a_i , $i = 1, 2, \dots, 5$ such that the zeros of $P(\lambda)$ have $\Re\lambda < 0$.

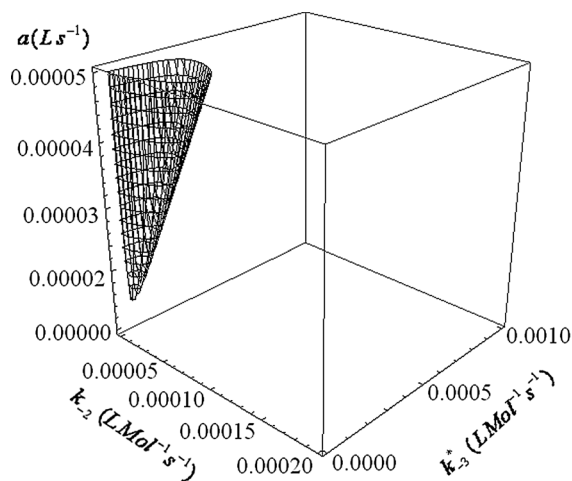


Fig. 4 Effect of varying the internal flow rate. The three dimensional figure represents all points for which the racemic state is *unstable* to perturbations. Increased flow enlarges the allowed region of instability. The racemic state is stable in the empty (white) domain. The remainder of parameters and initial concentrations as in Fig. 2.

The necessary and sufficient conditions for this to hold are the Routh-Hurwitz conditions. One such form, together with

$$a_5 \equiv -\det(B) \det(A) > 0, \quad (34)$$

is that

$$D_1 = a_1 \equiv -[\text{tr}(B) + \text{tr}(A)] > 0, \quad \& \quad (35)$$

$$D_2 = (a_1 a_2 - a_3) > 0, \quad \& \quad (36)$$

$$D_3 = a_3 D_2 + a_1 (a_5 - a_1 a_4) > 0, \quad \& \quad (37)$$

$$D_4 = a_4 D_3 - a_5 \{a_1 (a_2^2 - a_4) - (a_3 a_2 - a_5)\} > 0. \quad (38)$$

The expressions a_i can be read off directly from comparing the polynomial in eqn (33) to $-P$ calculated in eqn (56).

If *any* of the above conditions, eqn (34)–(38) does not hold, then the racemic fixed point solution is unstable. As before, this means the system will evolve to a chiral state. This test can be compared with direct numerical calculation of the five eigenvalues λ_i , $i = 1, 2, 3, 4, 5$ (the roots of the characteristic polynomial). We find complete agreement between the two methods (eigenvalues, RH criteria) employed. We emphasize that an instability in the racemic fixed point implies the onset of a bifurcation to a non-racemic, hence chiral, solution.

4.2 Domains of instability

We initiate the procedure outlined above by specifying the forward/reverse reaction rates (temperature differences are treated implicitly), the internal flow rate, the compartment volumes and the conserved total system mass:

$$\{k_{\pm i}(T), k_{\pm i}^*(T^*), a, V, V^*, C, \quad (39)$$

the individual rates of course satisfying microreversibility at the respective temperatures T and T^* . We then solve for the

complete racemic fixed point solution, retaining only those solutions that are non-negative:

$$\left\{ \frac{dA}{dt} = 0, \frac{dZ}{dt} = 0, \frac{dZ^*}{dt} = 0 \right\}_{\bar{y}=\bar{y}^*=0} \Rightarrow \{\bar{A}, \bar{Z}, \bar{Z}^*\}_{\geq 0}. \quad (40)$$

We next evaluate the Jacobian matrix over this fixed point solution:

$$J \Rightarrow \left(\begin{array}{cc} A^{3 \times 3} & 0 \\ 0 & B^{2 \times 2} \end{array} \right) \Big|_{\{\bar{A}, \bar{Z}, \bar{Z}^*\}_{\geq 0}}. \quad (41)$$

As a final step we evaluate the five RH conditions eqn (34)–(38): are they *all* true or not? If not, then we immediately know that the racemic fixed point is unstable for the parameter choice made in eqn (39). Hence any fluctuation about the idealized racemic composition will grow and drive the system to a chiral final state. In parallel, we also evaluate numerically the five roots of the characteristic polynomial λ_i , $i = 1, 2, 3, 4, 5$ and verify agreement between RH criteria and the eigenvalues. We use the RH criteria to map out the regions of linear instability for the racemic solution.

We start with the reaction rates $k_1 = 10^{-9}$, $k_{-1} = 10^{-15}$, $k_1^* = 10^{-7}$, $k_{-1}^* = 10^{-11}$, $k_2 = 10^2$, $k_{-2} = 10^{-4}$, $k_2^* = 0$, $k_{-2}^* = 0$, $k_3 = 0$, $k_{-3} = 0$, $k_3^* = 10^1$, $k_{-3}^* = 10^{-3}$, the volume and flux parameters $V = 10$, $V^* = 10$, $a = 0.1$, and the initial concentrations $[A]_0 = 10^{-6}$, $[A^*]_0 = 10^{-6}$, $[L]_0 = 10^{-11}$, $[D]_0 = 10^{-11}$, $[L^*]_0 = 10^{-11} + 10^{-21}$, $[D^*]_0 = 10^{-11}$ as employed in ref. 9. This point in parameter space was shown to lead to SMSB and subsequent chiral amplification from direct numerical integration of the differential rate equations. These specific rate values and system parameters were obtained after performing a set of trial and error numerical simulations obeying the conditions $k_{-3}^* > k_{-2}$ and $k_2 > k_3^*$ which correspond approximately to thermodynamically unattainable conditions for systems with a uniform temperature and lacking compartments. Once obtained, we can exploit the stability analysis to map out and amplify the full domain of instability of the racemic solution.

We consider how the stability of the racemic fixed point responds to variations of selected pairs of variables about this point. In Fig. 2 we display the regions of stability/instability as a function of (a) the two compartment volumes V and V^* , (b) the flow rate a versus equal compartment volumes $V = V^*$, (c) the total concentration C versus flow rate a , and (d) $\alpha = \frac{k_{-3}^*}{k_{-2}}$ versus k_{-2} . The graphs in (a) and (b) merely tell us that the ratio of the compartment volumes cannot be either arbitrarily large nor small, and that for equal compartment volumes, there is a minimum flow rate below which no chiral state can be obtained. Perhaps more surprising are the trends indicated in Fig. 2(c) and (d): for a given flow rate there is a critical system concentration *above* which mirror symmetry cannot be broken and in (d), for a given value of k_{-2} a bounded region in k_{-3}^* in which the racemic state is unstable, and this region narrows down and pinches off for a critical value of k_{-2} . We can here appreciate that the conjecture $k_{-3}^* > k_{-2}$ is necessary, but certainly not sufficient to lead to an instability in the racemic state.

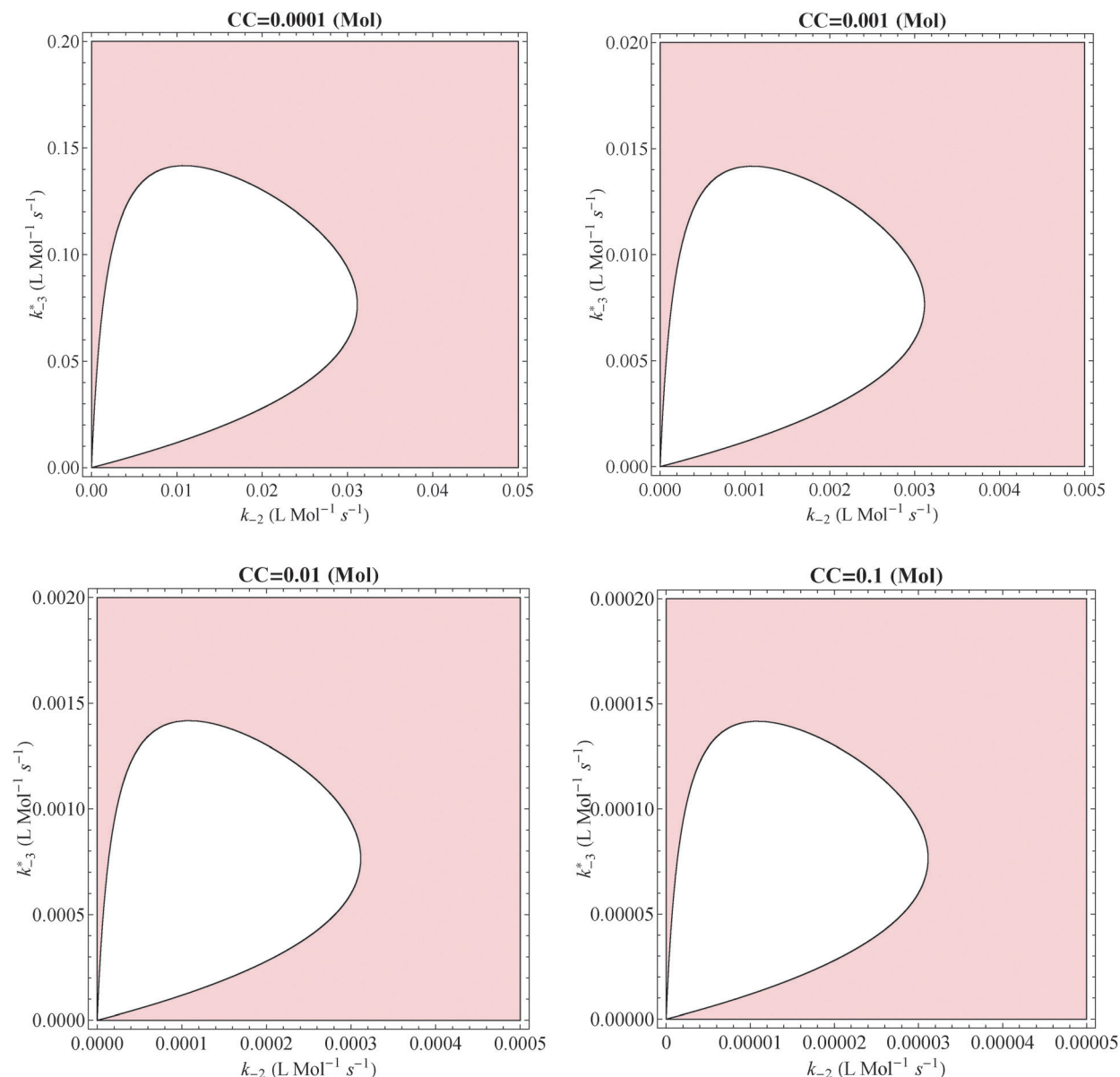


Fig. 5 Regions where the racemic state bifurcates to chiral states (in white) for different values of k_{-2} and k_{-3}^* , and for different values of the total concentration C . Increasing C decreases the ranges in both k_{-2} and k_{-3}^* for which the racemic state is linearly unstable. The remainder of parameters and initial concentrations as in Fig. 2.

We therefore delve further into this nonlinear relationship by considering how the instability varies with k_{-3}^* , k_{-2} and with the flow rate a . In Fig. 3 we display a sequence of plots showing how the region of instability grows in area as we scale up the flow rate. For a given flow rate, there is a maximum value of k_{-2} beyond which no value of k_{-3}^* will lead to an instability, and this maximum value scales with a . Below this maximum value, there are always upper and lower bounds on k_{-3}^* between which the system is unstable. These bounds also scale with a . In Fig. 3 we keep both k_{-3}^* and k_{-2} fixed and vary a : increased flow enlarges the region of instability. This can also be appreciated in the three-dimensional plot in Fig. 4 exhibiting how the region of instability expands in the cross-sectional area as we scale up the internal flow rate a .

In Fig. 5 we display a sequence of plots showing how the region of instability decreases in area as we scale up the total concentration C , over three orders of magnitude. As C is scaled up, the rates k_{-3}^* , k_{-2} decrease in magnitude in such a way as to preserve the *shape* of the region of instability. As before, for a given total concentration, there is a maximum value of k_{-2} beyond which no value of k_{-3}^* will lead to an instability, and this maximum value inversely scales with C . Below this maximum value, there are always upper and lower bounds on k_{-3}^* between which the system is unstable. These bounds also inversely scale with C . A three-dimensional plot in Fig. 6 indicates the cross-sectional area of the domain of instability shrinks as we scale up the total system concentration C . Dilute concentrations are more favorable for SMSB.⁹

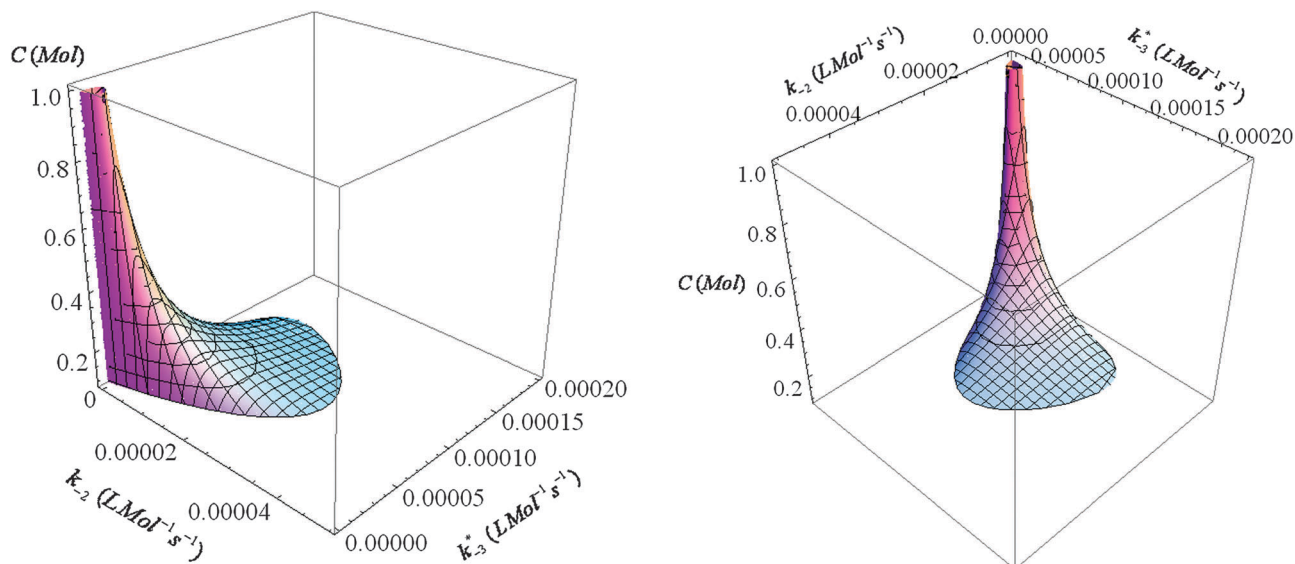


Fig. 6 Effect of varying the total system concentration C . The solid three-dimensional figure (color/gray) indicates where the racemic state is *unstable* to perturbations. Increasing the system concentration C *shrinks* the cross-sectional domain k_{-2}, k_{-3} of instability. The racemic state is linearly stable in the empty (white) region. The remainder of parameters and initial concentrations are as listed in Fig. 2.

5 Symmetry breaking and entropy production

We illustrate the above considerations by way of two examples. The inherent chiral fluctuations about the ideal racemic composition can be modeled by starting with an initial *ee* below the statistical deviation. In the first simulation, Fig. 7, we begin with $[A]_0 = [A^*]_0 = 1 \times 10^{-6}$ M, $[L]_0 = [D]_0 = 1 \times 10^{-11}$ M, $[L^*]_0 = 1 \times 10^{-6} + 1 \times 10^{-21}$ M, and $[D^*]_0 = 1 \times 10^{-6}$ M. In the second, Fig. 8, we keep the same rates and system parameters but start off with different initial concentrations: $[A]_0 = [A^*]_0 = 1 \times 10^{-11}$ M, $[L]_0 = [D]_0 = 5 \times 10^{-7}$ M, $[L^*]_0 = 5 \times 10^{-7} + 1 \times 10^{-20}$ M, and $[D^*]_0 = 5 \times 10^{-7}$ M.

We calculate and superimpose the entropy production²¹ σ on the concentration curves. The entropy production is a measure of the dissipation of the out-of-equilibrium process under study. For the initial conditions leading to Fig. 7, the entropy production suffers a peak at the onset of the induction

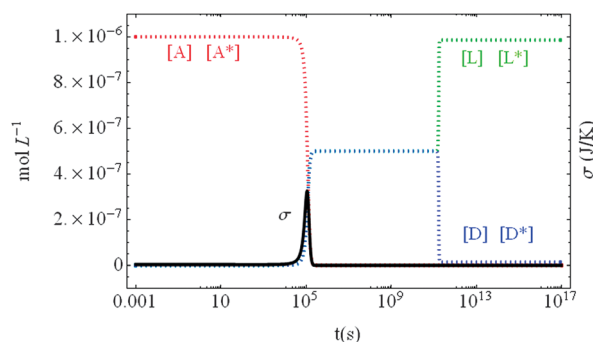


Fig. 7 Symmetry breaking bifurcation in the two compartment model, see Fig. 1. The rates and system parameters are given in Section 4.2. The entropy production σ peaks at the onset of the induction period, well before the symmetry breaking bifurcation. Note $\sigma > 0$ for all times.

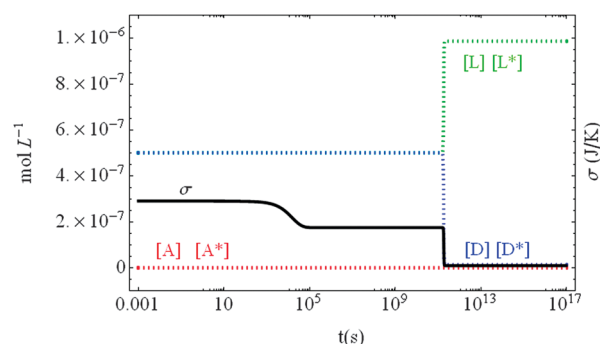


Fig. 8 Symmetry breaking bifurcation in the two compartment model, same rates and system parameters as in Fig. 7. See the text for the initial concentrations. In this case the entropy production σ (scaled up by a factor of 10^7) decreases monotonically and drops to a minimum (but nonzero) value at the symmetry breaking bifurcation, and remains positive for all subsequent times.

period, reflecting the fact that the catalysis of enantiomers, eqn (2), and subsequent depletion of the achiral substrate, is the most important contribution. The production then falls to a small but constant nonzero value and remains nonzero as long as the system is kept out of equilibrium. In the case of Fig. 8, the entropy production starts off from a large value then drops (at the same time scale as the peak in Fig. 7) and subsequently drops down to its minimum nonzero value at the symmetry breaking bifurcation. It again remains small but nonzero as long as the system is kept out of equilibrium.

6 Concluding remarks

Compartmentalization, together with different temperatures for the enantioselective and the non-enantioselective autocatalyses, was shown to be a necessary condition for SMSB.

The temperature gradient and internal flow or recycling of hot and cold material are the required driving forces to keep the system far from equilibrium. On the basis of the stability analysis of the racemic final state, the results presented here demonstrate the existence of SMSB as a closed region (see Fig. 3, 5 and 6) in the phase representation of the reaction parameters. Explicit examples of SMSB are displayed in Fig. 7 and 8 showing also the time dependence of the associated production of entropy for starting from two different initial concentrations.

The considerations given to the microreversibility condition in Sections 1–3 serve to underscore the fact that limited enantioselectivity – of and by itself – cannot lead to spontaneous mirror symmetry breaking, neither for a uniform temperature (*i.e.* under an experimental condition required to attain chemical equilibrium), nor even for two temperatures (*i.e.* an experimental condition that excludes chemical equilibrium). By contrast, mirror symmetry breaking may be possible when in addition to a non-uniform temperature distribution the reactions (2) and (3) occur in two distinct compartments held at different temperatures. The microreversibility arguments,^{8,22} used as a proof of the correctness of the model, are necessary because we assume a scenario where chemical kinetics can be applied.

We briefly point out what modifications would have to be made in order to consider spatial temperature gradients. First, we need to input the temperature profile $T(x,y,z)$ as a known field. This converts the reaction rates k_i into spatially dependent functions *via* the Arrhenius relation. The kinetic rate equations must be replaced by partial differential reaction-diffusion-advection equations for the spatial and temporal dependence of the concentrations. The background convective flow of hot/cold material could be modeled by a steady hydrodynamical flow-field compatible with the compartment boundary conditions. In short, the technical complications would be considerable, but not insurmountable.

The significance of the above results with respect to applied absolute synthesis is somewhat limited due to the temperature difference required for achieving the essential inequality $k_{-3}^* > k_{-2}$. The point is, a sufficiently large difference would require high pressure conditions as well to maintain the media in a liquid state. By contrast, the results can be of significance in scenarios of prebiotic chemistry. In this context, the required experimental conditions for high temperature gradients and compartmentalization agree with those found in deep ocean hydrothermal vents, as we have previously reported.⁹ Furthermore, the more relevant feature of this LES scenario is its ability to exhibit SMSB at very low reactant concentrations. In this respect, the detailed description given here for the effect on the SMSB dependence of the minimal internal flow rate – with respect to the compartment volumes at different temperatures – and on the reaction rate inequality $k_{-3}^* > k_{-2}$ shows that the variability range for SMSB of these parameters increases when the total system concentration C decreases, *i.e.* the probabilities for SMSB to occur *increase* in a prebiotic scenario. Note that this overcomes the more important difficulty for a reasonable SMSB in a prebiotic scenario, where a specific prebiotic organic compound could be present, but only at very low concentrations. This, in spite of the presence of an important fraction of organic

compounds, is because of the diversity of different organic compounds in prebiotic scenarios.²³ While LES is ruled out as a scheme for SMSB under experimental conditions where thermodynamic equilibrium can be achieved, the results presented here open up the study for SMSB in scenarios of non-uniform temperature distributions, specific energy inputs to some species of the system, and compartmentalization, *i.e.* under conditions similar to those of living systems.

Appendix

A Fluctuation equations

The linearized fluctuation equations that follow from the kinetic equations, eqn (7)–(12) in Section 2.1 are as follows. The overbar denotes a stationary solution of eqn (7)–(12).

$$\begin{aligned} \delta \dot{A} = & \left(-2k_1 - [k_2 + k_3]\bar{\chi} - \frac{a}{V} - \frac{a}{V^*} \right) \delta A \\ & + \left(-(k_2 + k_3)\bar{A} + (k_{-2} + k_{-3})\bar{\chi} + k_{-1} - \frac{a}{V^*} \right) \\ & \times \delta \chi + (k_{-2} - k_{-3})\bar{y} \delta y - \frac{a}{V} \delta \chi^* \end{aligned} \quad (42)$$

$$\begin{aligned} \delta \dot{\chi} = & (2k_1 + [k_2 + k_3]\bar{\chi}) \delta A \\ & + \left((k_2 + k_3)\bar{A} - k_{-1} - (k_{-2} + k_{-3})\bar{\chi} - \frac{a}{V} \right) \delta \chi \\ & + (k_{-3} - k_{-2})\bar{y} \delta y + \frac{a}{V} \delta \chi^* \end{aligned} \quad (43)$$

$$\begin{aligned} \delta \dot{y}^* = & (k_2 - k_3)\bar{y} \delta A - k_{-2}\bar{y} \delta \chi \\ & + \left(-k_{-1} - k_{-2}\bar{\chi} + (k_2 - k_3)\bar{A} - \frac{a}{V} \right) \delta y + \frac{a}{V^*} \delta y^* \end{aligned} \quad (44)$$

$$\begin{aligned} \delta \dot{\chi}^* = & -\frac{V}{V^*} (2k_1^* + [k_2^* + k_3^*]\bar{\chi}^*) \delta A \\ & + \left(\frac{a}{V^*} - \frac{V}{V^*} (2k_1^* + [k_2^* + k_3^*]\bar{\chi}^*) \right) \delta \chi \\ & + \left([k_2^* + k_3^*] \left(\frac{C}{V^*} - \frac{V}{V^*} (\bar{A} + \bar{\chi}) - \bar{\chi}^* \right) \right. \\ & \left. - k_{-1}^* - (k_{-2}^* + k_{-3}^*)\bar{\chi}^* - \frac{a}{V^*} \right. \\ & \left. - (2k_1^* + [k_2^* + k_3^*]\bar{\chi}^*) \right) \delta \chi^* + (k_{-3}^* - k_{-2}^*)\bar{y}^* \delta y^* \end{aligned} \quad (45)$$

$$\begin{aligned} \delta \dot{y}^* = & -\frac{V}{V^*} \bar{y}^* (k_2^* - k_3^*) \delta A - \frac{V}{V^*} \bar{y}^* (k_2^* - k_3^*) \delta \chi + \frac{a}{V^*} \delta y \\ & - (k_{-2}^* + k_2^* - k_3^*)\bar{y}^* \delta \chi^* \\ & + \left((k_2^* - k_3^*) \left(\frac{C}{V^*} - \frac{V}{V^*} (\bar{A} + \bar{\chi}) - \bar{\chi}^* \right) \right. \\ & \left. - k_{-1}^* - k_{-2}^* \bar{\chi}^* - \frac{a}{V^*} \right) \delta y^* \end{aligned} \quad (46)$$

Specializing to the racemic fixed point $\bar{y} = \bar{y}^* = 0$ leads to the decoupling of the first three fluctuations from the latter two:

$$\begin{aligned} \delta \dot{A} = & \left(-2k_1 - [k_2 + k_3]\bar{\chi} - \frac{a}{V} - \frac{a}{V^*} \right) \delta A \\ & + \left(-(k_2 + k_3)\bar{A} + (k_{-2} + k_{-3})\bar{\chi} + k_{-1} - \frac{a}{V^*} \right) \delta \chi - \frac{a}{V} \delta \chi^* \end{aligned} \quad (47)$$

$$\begin{aligned} \delta \dot{\chi} = & (2k_1 + [k_2 + k_3]\bar{\chi}) \delta A \\ & + \left((k_2 + k_3)\bar{A} - k_{-1} - (k_{-2} + k_{-3})\bar{\chi} - \frac{a}{V} \right) \delta \chi + \frac{a}{V} \delta \chi^* \end{aligned} \quad (48)$$

$$\delta \dot{y} = + \left(-k_{-1} - k_{-2}\bar{\chi} + (k_2 - k_3)\bar{A} - \frac{a}{V} \right) \delta y + \frac{a}{V} \delta y^* \quad (49)$$

$$\begin{aligned} \delta \dot{\chi}^* = & -\frac{V}{V^*} (2k_1^* + [k_2^* + k_3^*]\bar{\chi}^*) \delta A \\ & + \left(\frac{a}{V^*} - \frac{V}{V^*} (2k_1^* + [k_2^* + k_3^*]\bar{\chi}^*) \right) \delta \chi \\ & + \left([k_2^* + k_3^*] \left(\frac{C}{V^*} - \frac{V}{V^*} (\bar{A} + \bar{\chi}) - \bar{\chi}^* \right) \right. \\ & \left. - k_{-1}^* - (k_{-2}^* + k_{-3}^*)\bar{\chi}^* - \frac{a}{V^*} \right. \\ & \left. - (2k_1^* + [k_2^* + k_3^*]\bar{\chi}^*) \right) \delta \chi^* \end{aligned} \quad (50)$$

$$\begin{aligned} \delta \dot{y}^* = & + \frac{a}{V^*} \delta y \\ & + \left((k_2^* - k_3^*) \left(\frac{C}{V^*} - \frac{V}{V^*} (\bar{A} + \bar{\chi}) - \bar{\chi}^* \right) - k_{-1}^* - k_{-2}^* \bar{\chi}^* - \frac{a}{V^*} \right) \delta y^*, \end{aligned} \quad (51)$$

as reflected by the specific matrix entries in eqn (31) and (32).

B Characteristic polynomial

The characteristic polynomial associated with the Jacobian matrix, eqn (29), evaluated at the racemic fixed point is

$$\begin{aligned} P(\lambda) = \det[J - \lambda I] &= \det[A - \lambda I] \det[B - \lambda I], \\ &= P^{(3)}(\lambda) P^{(2)}(\lambda), \end{aligned} \quad (52)$$

where I is the identity matrix and the quadratic and cubic polynomials are given by

$$P^{(2)}(\lambda) = \lambda^2 - \text{tr}(B)\lambda + \det(B), \quad (53)$$

$$P^{(3)}(\lambda) = -\lambda^3 + \text{tr}(A)\lambda^2 + G(A)\lambda + \det(A), \quad (54)$$

respectively, where

$$G(A) = -a_{11}a_{22} - a_{11}a_{33} - a_{22}a_{33} + a_{32}a_{23} + a_{12}a_{21} + a_{13}a_{31}. \quad (55)$$

Thus, inserting eqn (53) and (54) into eqn (52), we obtain the fifth order polynomial

$$\begin{aligned} P(\lambda) = & -\lambda^5 + [\text{tr}(B) + \text{tr}(A)]\lambda^4 + \{G(A) - \det(B) \\ & - \text{tr}(B)\text{tr}(A)\}\lambda^3 + \{\text{tr}(A)\det(B) - G(A)\text{tr}(B) + \det(A)\}\lambda^2 \\ & + \{G(A)\det(B) - \det(A)\text{tr}(B)\}\lambda + \det(B)\det(A). \end{aligned} \quad (56)$$

Acknowledgements

The research of CB and DH is supported in part by the Grant AYA2009-13920-C02-01, and that of JMR, JC, ZE-H and AM by AYA2009-13920-C02-02, from MICINN (currently MINECO).

References

- P. Cintas and C. Viedma, *Chem. Commun.*, 2011, **47**, 12786.
- Z. El-Hachemi, J. Crusats, J. Ribó, J. McBride and S. Veintemillas-Verdaguer, *Angew. Chem., Int. Ed.*, 2011, **50**, 2359.
- W. Noorduyn, T. Izumi, A. Millemaggi, M. Leeman, H. Meekes, W. van Enkevort, R. Kellogg, B. Kaptein, E. Vlieg and D. Blackmond, *J. Am. Chem. Soc.*, 2008, **130**, 1158.
- W. Noorduyn, E. Vlieg, R. Kellogg and B. Kaptein, *Angew. Chem., Int. Ed.*, 2009, **48**, 9600.
- C. Viedma, *Phys. Rev. Lett.*, 2005, **94**, 065504.
- J. Wattis, *Origins Life Evol. Biospheres*, 2011, **41**, 133.
- V. Avetisov and V. Goldanskii, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 11435.
- J. Ribó and D. Hochberg, *Phys. Lett. A*, 2008, **373**, 111.
- J. Ribó, J. Crusats, Z. El-Hachemi, A. Moyano, C. Blanco and D. Hochberg, *Astrobiology*, accepted, AST-2012-0904.R1.
- J. Baross and S. Hoffman, *Origins Life Evol. Biospheres*, 1985, **15**, 327.
- N. Holm, *Origins Life Evol. Biospheres*, 1992, **22**, 5.
- W. Martin, J. Baross, D. Kelley and M. Russell, *Nat. Rev. Microbiol.*, 2008, **6**, 805.
- S. Miller and J. Bada, *Nature*, 1988, **334**, 609.
- J. Bernal, *The Physical Basis of Life*, Routledge, London, 1951.
- A. Brack, *Clay minerals and the origin of life*, in *Developments in Clay Science Handbook of Clay Science*, Elsevier, Oxford, 2006, vol. 1.
- A. Cairns-Smith and H. Hartman, *Clay minerals and the origin of life*, Cambridge University Press, UK, 1986.
- J. Ferris, *Elements*, 2005, **1**, 145.
- J. Ferris, *Am. Mineral.*, 2006, **91**, 1715.
- R. Chang, *Physical Chemistry*, University Science Books, Sausalito, 2000.
- J. Murray, *Mathematical Biology*, Springer, Sausalito, 2nd edn, 1993.
- D. Kondepudi and I. Prigogine, *Modern Thermodynamics: From heat engines to dissipative structures*, Wiley, New York, Sausalito, 2nd edn, 1998.
- D. Blackmond and O. Matar, *J. Phys. Chem. B*, 2008, **112**, 5098.
- P. Schmitt-Kopplina, Z. Gabelicab, R. D. Gougeonc, A. Feketea, B. Kanawatia, M. Harira, I. Gebefuegia, G. Eckeld and N. Hertkorna, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **105**, 2763.

Chiral polymerization: symmetry breaking and entropy production in closed systems

Celia Blanco* and David Hochberg*

Received 25th June 2010, Accepted 11th October 2010

DOI: 10.1039/c0cp00992j

We solve numerically a kinetic model of chiral polymerization in systems closed to matter and energy flow, paying special attention to its ability to amplify the small initial enantiomeric excesses due to the internal and unavoidable statistical fluctuations. The reaction steps are assumed to be reversible, implying a thermodynamic constraint among some of the rate constants. Absolute asymmetric synthesis is achieved in this scheme. The system can persist for long times in quasi-stationary chiral asymmetric states before racemizing. Strong inhibition leads to long-period chiral oscillations in the enantiomeric excesses of the longest homopolymer chains. We also calculate the entropy production σ per unit volume and show that σ increases to a peak value either before or in the vicinity of the chiral symmetry breaking transition.

I. Introduction

There is a growing consensus that the homochirality of biological compounds is a condition associated to life that emerged in the abiotic stages of evolution through processes of spontaneous mirror symmetry breaking (SMSB). This could have proceeded in a prebiotic stage, incorporating steps of increasing complexity thus leading to chemical systems and enantioselective chemical networks.^{1–3} An important issue is therefore to identify processes of chirality amplification in chemical reactions. In this regard, a recent kinetic analysis of the Frank model in closed systems applied to the Soai reaction⁴ has taught us that in an actual chemical scenario, reaction networks that exhibit SMSB are very sensitive to chiral inductions owing to the presence of tiny initial enantiomeric excesses, as previously shown theoretically.⁵ The stochastic scenario implies the creation of chirality from intrinsic chiral fluctuations and its later transmission and amplification. This can occur in far-from-equilibrium systems that undergo dynamic phase transitions.

The process must be coupled to others which preserve, extend, and transmit the chirality. Biological homochirality of living systems involves large macromolecules, therefore a central point is the relationship of the polymerization process with the emergence of chirality. This hypothesis has inspired recent activity devoted to modeling efforts aimed at understanding mirror symmetry breaking in polymerization of relevance to the origin of life. The models so proposed^{6–14} are by and large, elaborate extensions and generalizations of Frank's original paradigmatic scheme.¹⁵ Heading this list, Sandars⁶ introduced a detailed polymerization process plus the basic elements of enantiomeric cross inhibition as well as a chiral feedback mechanism in which only the largest polymers formed can enhance the production of the monomers from an achiral substrate. He treated basic numerical studies of

symmetry breaking and bifurcation properties of this model for various values of the number of repeat units N . All the subsequent models cited here are variations on Sandars' original theme. Soon afterwards, Brandenburg and coworkers⁸ studied the stability and conservation properties of a modified Sandars' model and introduce a reduced $N = 2$ version including the effects of chiral bias. In ref. 7, they included spatial extent in this model to study the spread and propagation of chiral domains as well as the influence of a background turbulent advection velocity field. The model of Wattis and Coveney⁹ differs from Sandars' in that they allow polymers to grow to arbitrary lengths N and the chiral polymers of all lengths, from the dimer and upwards, act catalytically in the breakdown of the achiral source into chiral monomers. An analytic linear stability analysis of both the racemic and chiral solutions is carried out for the model's large N limit and various kinetic timescales are identified. The role of external white noise on Sandars-type polymerization networks including spatial extent has been explored by Gleiser and coworkers: the $N = 2$ truncated model introduced in ref. 7 is subjected to external white noise in ref. 10, chiral bias is considered in ref. 11, high intensity and long duration noise is considered in ref. 12 and in ref. 13, modified Sandars-type models with spatial extent and external noise are considered both for finite and infinite N , with an emphasis paid to the dynamics of chiral symmetry breaking. By contrast, Saito and Hyuga's¹⁴ model gives rise to homochiral states but differs markedly from Sandars' in that it does not invoke the enantiomeric cross inhibition, allowing instead for reversibility in all the reaction steps. Their model requires open flow, which is the needed element of irreversibility. A different model which stands apart from the above group is that of Plasson *et al.*¹⁶ They considered a recycled system based on reversible chemical reactions and open only to energy flow and without any (auto)catalytic reactions. A source of constant external energy—the element of irreversibility—is required to activate the monomers. This energy could be introduced into the system in a physical form, say, as high energy photons. A system of this kind, limited to dimerizations, was shown to

Centro de Astrobiología (CSIC-INTA), Carretera Ajalvir Kilómetro 4, 28850 Torrejón de Ardoz, Madrid, Spain. E-mail: blancodtc@inta.es, hochbergd@inta.es

have nonracemic stable final states for various ranges of the model parameter values and for total concentrations greater than a minimal value.

The polymerization models referred to above are defined only for *open flow* systems which exchange matter and energy with the exterior. A constant source of achiral precursor is usually assumed. An unrealistic consequence is that homochiral chains can grow to infinite length. By contrast, most experimental procedures are carried out in *closed and spatially bounded* reaction domains and are initiated in far-from-equilibrium states.^{3,17–24} It is thus crucial to have models compatible with these experimentally realistic boundary and initial conditions. The most immediate consequences are that polymer chains can grow to a finite maximum length and that the total system mass is constant. These two properties are of course intimately related. An original aspect of our work is that we consider the polymerization process in a closed system and include reversible reactions. This enables us to explore absolute asymmetric synthesis in thermodynamically closed systems (closed to matter flow) taking into account the backward reaction steps, which we call here “reversible reactions/reversible models” in spite of the fact that the values of the forward and reverse rate constants can be very different. As we are eloquently reminded by Mislow, mirror asymmetric states are in practice unavoidable on purely statistical grounds alone,²⁵ even in the absence of chiral physical forces. Absolute asymmetric synthesis is the ability of a system to amplify these statistical and minuscule enantiomeric biases up to observably large excesses. Thus a major goal of this work is to assess the ability of such polymerization schemes to amplify these initial excesses, albeit if only temporarily. Asymmetric amplification is demonstrated to occur obeying microscopic reversibility in a reversible model closed to matter flow. This is of great practical interest as the racemization time scale can be substantially longer than that of the transformation of the initial reagents, depending on the strength of the mutual inhibition, or direct interaction between the enantiomers. And regarding chemical evolution, this obviates the need to invoke chiral physical fields and lends further support to the conviction that homochirality is a “stereochemical imperative” of molecular evolution.²⁶

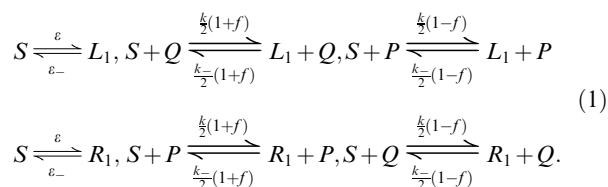
The behavior of entropy in polymerization models is rarely discussed,²⁷ and has not been addressed for mirror symmetry breaking in chiral polymerization. The entropy produced in a chemical reaction initiated out of equilibrium gives a measure of the dissipation during the approach to final equilibrium. In this paper, we calculate the rate of entropy production in chiral polymerization. Depending on the enantiomeric mutual inhibition, the entropy production undergoes a rapid increase to a peak value either before, or else in the vicinity of the chiral symmetry breaking transition, followed by an equally dramatic decrease. The system racemizes at time scales greater than that of polymerization, and is accompanied by a final decrease of entropy production to zero, indicating that the system has reached a final stable state (not to be confused with a stationary state, which must be accompanied by a nonzero constant entropy production). Computation of the average homopolymer length indicates

that the final racemic state is dominated by the longest available chains.

II. The polymerization model

The model we introduce and study here is modified and extended from that of Wattis and Coveney⁹ which is in turn, a generalization of Sandars’s scheme.⁶ Two salient differences that distinguish our model from these and other previous ones^{6–14} are that we (1) consider polymerization in *closed* systems²⁸—so that no matter flow is permitted with an external environment—and (2) we allow for reversible reactions in all the steps. A third difference is that we also include the formation (and dissociation) of the heterodimer. While heterodimer formation was originally contemplated in ref. 6, it has been silently omitted from all the subsequent models^{7–13} that derive therefrom. Fragmentation in a Sandars type model has been considered previously, but was shown to yield a maximum average polymer length of only $N = 3$ repeat units.²⁹

We assume there is an achiral precursor S which can directly produce the chiral monomers L_1 and R_1 at a slow rate ε as well as be consumed in processes in which homopolymers of all lengths catalyze the production of monomers. The specific reaction scheme we study here is composed of the following steps, where the ε , (ε_-) , k (k_-), *etc.*, denote the forward (reverse) reaction rate constants and $0 \leq f \leq 1$ is the fidelity of the feedback mechanism:



Here

$$Q = \sum_{n=1}^N nL_n, P = \sum_{n=1}^N nR_n, \tag{2}$$

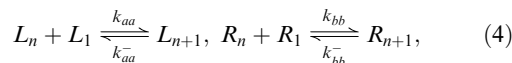
represents a measure of the total concentrations of left-handed and right handed polymers. The amount of each chiral monomer produced is influenced by the total amount of chiral oligomer already present in the system. We allow for the monomers themselves to participate in these substrate reactions: hence $N \geq n \geq 1$, where N is the maximum chain length. The central top and bottom reactions in eqn (1) are enantioselective, whereas those on the right hand side are non-enantioselective. The model therefore contains the features of limited enantioselectivity, first proposed as an alternative to the mutual inhibition of Frank.¹

An important observation is that differences in the Gibbs free energy ΔG^0 between initial and final states should be the same in all the reactions listed in eqn (1), which implies the thermodynamic constraint on the following forward and reverse reaction constants (see also ref. 14)

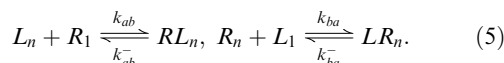
$$\frac{\varepsilon}{\varepsilon_-} = \frac{k}{k_-}. \tag{3}$$

The polymerization and chain end-termination reactions (see below) are not subject to a thermodynamic constraint.

The monomers combine to form chirally pure polymer chains denoted by L_n and R_n , according to the isodesmic³¹ stepwise reactions for $1 \leq n \leq N-1$

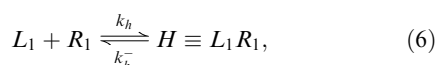


and inhibition, or the chain end-termination reactions for $N-1 \geq n \geq 2$:



These upper limits for n specified in eqn (4) and (5) ensure that the *maximum* length for all oligomers produced (or consumed) by these reaction sets, both the homo- and heterochiral ones, is never greater than N . In the remainder of this paper we will consider here the natural and chiral symmetric rate assignments $k_{aa} = k_{bb}$, $k_{ab} = k_{ba}$ and likewise for the inverse rates, $k_{aa}^- = k_{bb}^-$ and $k_{ab}^- = k_{ba}^-$. There are then four independent rate constants.

We include a separate reaction for the heterodimer formation and dissociation:



where $k_h = (k_{ab} + k_{ba})/2$ and $k_h^- = (k_{ab}^- + k_{ba}^-)/2$. Note that these latter two rate constants are automatically determined from the above choice and that $L_1 R_1 = R_1 L_1$ of course. This completes the specification of the model's reactions.

The model is left-right symmetric, that is, possesses a discrete Z_2 symmetry,³⁰ which is manifest in the elementary reaction steps, in the rate constants, and in the corresponding differential rate equations (see below). This symmetry can be broken spontaneously by the dynamical solutions of the differential rate equations. This model is thus apt for investigating spontaneous mirror symmetry breaking. Though not considered here, the effects of explicit chiral bias (*e.g.*, that induced by external physical fields) can also be studied with this model by lifting the Z_2 degeneracy in the reaction rates, *e.g.*, allowing for $k_{aa} \neq k_{bb}$, *etc.*, leading to a maximum of eight independent rate constants for describing the reaction set in eqn (4) and (5).

Rate-equation theory as employed in chemical kinetics is used to describe the differential rate equations of the achiral source, the monomers, as well as the homo- and heterochiral oligomers belonging to this reaction network. The kinetic equations for the concentrations that follow from these elementary steps are as follows. We begin with the rate equations for the achiral precursor S and those

corresponding to the two chiral monomers:

$$\begin{aligned} \frac{d[S]}{dt} = & -2\varepsilon[S] - k[S](P+Q) + \varepsilon_-([L_1] + [R_1]) \\ & + \frac{1}{2}k_-[L_1]((1+f)Q + (1-f)P) \\ & + \frac{1}{2}k_-[R_1]((1+f)P + (1-f)Q), \end{aligned} \quad (7)$$

$$\begin{aligned} \frac{d[L_1]}{dt} = & \varepsilon[S] + \frac{k}{2}[S]((1+f)Q + (1-f)P) \\ & - k_{aa}[L_1] \left(2[L_1] + \sum_{n=2}^{N-1} [L_n] + k_{ba}/k_{aa} \sum_{n=1}^{N-1} [R_n] \right) \\ & - \varepsilon_-[L_1] - \frac{k_-}{2}[L_1]((1+f)Q + (1-f)P) \\ & + k_{aa}^- \left(2[L_2] + \sum_{n=2}^{N-1} [L_{n+1}] + k_{ba}^-/k_{aa}^- \sum_{n=2}^{N-1} [LR_n] \right) + k_h^- H, \end{aligned} \quad (8)$$

$$\begin{aligned} \frac{d[R_1]}{dt} = & \varepsilon[S] + \frac{k}{2}[S]((1+f)P + (1-f)Q) \\ & - k_{bb}[R_1] \left(2[R_1] + \sum_{n=2}^{N-1} [R_n] + k_{ab}/k_{bb} \sum_{n=1}^{N-1} [L_n] \right) \\ & - \varepsilon_-[R_1] - \frac{k_-}{2}[R_1]((1+f)P + (1-f)Q) \\ & + k_{bb}^- \left(2[R_2] + \sum_{n=2}^{N-1} [R_{n+1}] + k_{ab}^-/k_{bb}^- \sum_{n=2}^{N-1} [RL_n] \right) + k_h^- H. \end{aligned} \quad (9)$$

Whereas for $N-1 \geq n \geq 2$ we have the set of stepwise polymerization rate equations

$$\begin{aligned} \frac{d[L_n]}{dt} = & k_{aa}[L_1]([L_{n-1}] - [L_n]) - k_{ab}[L_n][R_1] \\ & + k_{ab}^-[RL_n] + k_{aa}^-([L_{n+1}] - [L_n]), \end{aligned} \quad (10)$$

$$\begin{aligned} \frac{d[R_n]}{dt} = & k_{bb}[R_1]([R_{n-1}] - [R_n]) - k_{ba}[R_n][L_1] \\ & + k_{ba}^-[LR_n] + k_{bb}^-([R_{n+1}] - [R_n]). \end{aligned} \quad (11)$$

Note, in accord with eqn (4) and (5) for the largest polymers $n = N$, we have instead the final pair

$$\frac{d[L_N]}{dt} = k_{aa}[L_1][L_{N-1}] - k_{aa}^-[L_N], \quad (12)$$

$$\frac{d[R_N]}{dt} = k_{bb}[R_1][R_{N-1}] - k_{bb}^-[R_N]. \quad (13)$$

Then the kinetic equation for the heterodimer $H \equiv L_1 R_1$ (which we keep separate from the other end-chain rate equations):

$$\frac{d[H]}{dt} = k_h[L_1][R_1] - k_h^-[H]. \quad (14)$$

Lastly, for $N - 1 \geq n \geq 2$ the rate equations for the “end-spoiled” chains:

$$\frac{d[LR_n]}{dt} = k_{ba}[L_1][R_n] - k_{ba}^- [LR_n], \quad (15)$$

$$\frac{d[RL_n]}{dt} = k_{ab}[R_1][L_n] - k_{ab}^- [RL_n]. \quad (16)$$

For chemical systems closed to matter flow, the constant mass constraint that must be obeyed by the coupled system of differential equations at every instant is given by (the overdot denotes time derivative):

$$\begin{aligned} [\dot{S}] + 2[\dot{H}] + \sum_{n=1}^N n([\dot{L}_n] + [\dot{R}_n]) \\ + \sum_{n=2}^{N-1} (n+1)([\dot{R}L_n] + [\dot{L}R_n]) = 0. \end{aligned} \quad (17)$$

This relation can be verified directly using the above set of $4N - 2$ kinetic equations eqn (7)–(16).

III. Entropy production

The entropy production rate in an irreversible process is a measure of the dissipation in that process. At equilibrium, the entropy production rate vanishes and is an extremum.³² This production has been investigated recently for reversible versions of the Frank model.^{33,34} In those simple models, the behavior of the entropy produced near the chiral symmetry breaking transition as well as its subsequent temporal development, depends on whether the chemical system is open or closed to matter flow. We will return to this important point below. Here we consider the behavior of the entropy produced by polymerization reactions and monomer catalysis when the system undergoes a chiral phase transition as well at the later stages when the system reaches final chemical equilibrium upon racemization.

For reactions obeying mass action kinetics, the entropy produced in any chemical reaction can be calculated straightforwardly in terms of the individual elementary reaction rates.^{32,35} The rate of entropy production is the sum over the difference of the forward (R_{jf}) and reverse (R_{jr}) reaction rates multiplied by the natural logarithm of the ratio of the forward and reverse rates:³⁵

$$\sigma(t) = R^* \sum_j (R_{jf} - R_{jr}) \ln \left(\frac{R_{jf}}{R_{jr}} \right) \geq 0, \quad (18)$$

where the sum runs over each elementary reaction step j , and $R^* = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$ is the gas constant. Since our reaction scheme is set up for closed systems, equilibrium is reached after a racemization time scale t_{racem} is reached, which suggests a further measure is provided by the total net entropy produced from the start of the reactions through chiral symmetry breaking, then on to final racemization, when the system reaches chemical and thermodynamic equilibrium and $\sigma(t > t_{\text{racem}}) = 0$

$$\sigma_T = \int_0^{t_{\text{racem}}} \sigma(u) du. \quad (19)$$

This quantifies the total dissipation over the complete history of the chemical transformations under study.

The sum in eqn (18) contains $2N + 4$ terms. In order to determine which specific steps of the full reaction network provide the leading contributions to σ , we group the forward and reverse reaction rates as follows:

A. Direct monomer production

$$R_{1f} = \varepsilon[S] \quad R_{1r} = \varepsilon_-[L_1], \quad (20)$$

$$R_{2f} = \varepsilon[S] \quad R_{2r} = \varepsilon_-[R_1]. \quad (21)$$

B. Monomer catalysis

$$R_{3f} = \frac{k}{2}(1+f)[S][Q] \quad R_{3r} = \frac{k_-}{2}(1+f)[L_1][Q], \quad (22)$$

$$R_{4f} = \frac{k}{2}(1+f)[S][P] \quad R_{4r} = \frac{k_-}{2}(1+f)[R_1][P], \quad (23)$$

$$R_{5f} = \frac{k}{2}(1-f)[S][P] \quad R_{5r} = \frac{k_-}{2}(1-f)[L_1][P], \quad (24)$$

$$R_{6f} = \frac{k}{2}(1-f)[S][Q] \quad R_{6r} = \frac{k_-}{2}(1-f)[R_1][Q]. \quad (25)$$

C. Polymerization

For $1 \leq n \leq N - 1$ we have

$$R_{nf}^L = k_{aa}[L_1][L_n] \quad R_{nr}^L = k_{aa}^- [L_{n+1}], \quad (26)$$

$$R_{nf}^R = k_{bb}[R_1][R_n] \quad R_{nr}^R = k_{bb}^- [R_{n+1}]. \quad (27)$$

D. End-chain termination and the heterodimer

The heterodimer rates are

$$R_f^h = k_h[L_1][R_1] \quad R_r^h = k_h^- [H], \quad (28)$$

whereas for $2 \leq n \leq N - 1$ the end-chain forward and reverse rates are given by

$$R_{nf}^{eR} = k_{ab}[L_n][R_1] \quad R_{nr}^{eR} = k_{ab}^- [RL_n], \quad (29)$$

$$R_{nf}^{eL} = k_{ba}[R_n][L_1] \quad R_{nr}^{eL} = k_{ba}^- [LR_n]. \quad (30)$$

IV. Results

As discussed in the Introduction, we are interested in testing the model’s ability to amplify the initial small statistical deviations about the idealized racemic composition,²⁵ in systems closed to matter flow and taking microscopic reversibility fully into account. In order to study the sensitivity of the above reversible polymerization network to these minuscule initial enantiomeric excesses, a very dilute initial concentration of a scalemic (non racemic) mixture was employed in the calculations: the initial monomeric concentrations of $[L_1]_0 = (1 \times 10^{-6} + 1 \times 10^{-15}) \text{ M}$ and

$[R_1]_0 = 1 \times 10^{-6}$ M yielding an initial chiral excess of $ee_0 = 5 \times 10^{-8}\%$. This is actually slightly lower^{25,36} than the excess corresponding to the initial monomer concentrations ($ee_0 = 6.1 \times 10^{-8}\%$). The initial concentration of the achiral substrate is $[S]_0 = 2$ M, whereas those corresponding the homo- hetero-oligomers are all set to zero: $[L_n]_0 = [R_n]_0 = 0$, for $2 \leq n \leq N$, $[H]_0 = 0$, and $[LR_n]_0 = [RL_n]_0 = 0$ for $2 \leq n \leq N-1$. We choose $\varepsilon = 2 \times 10^{-5}$, $\varepsilon_- = 10^{-10}$, $k = 2.0$, $k_- = 10^{-5}$, $f = 0.9$, $k_{aa} = k_{bb} = 1.0$, $k_{ab} = k_{ba} = 10^3$, $k_{ab}^- = k_{ba}^- = 1.0$, $k_{aa}^- = k_{bb}^- = 10^{-5}$. For illustrative purposes, we consider chains that can grow to a maximum length of $N = 12$. This N value is intended as a mean “ball-park” figure suggested by recent experiments yielding homochiral chains anywhere from $N = 5^{18}$ up to $N = 18^{24}$ chiral repeat units, depending on the amino acids employed and the experimental conditions. The differential rate equations eqn (7)–(16) were numerically integrated with the version 7 Mathematica program package and using a level of numerical precision, typically twenty or more significant digits, to ensure the numerical significance of these initial concentrations and enantiomeric excess. The results were monitored and verified to assure that total system mass eqn (17) remained constant in time.

Results are quantified in terms of a variety of convenient chiral measures. The percent enantiomeric excess values of the oligomers with homochiral sequences are calculated according to ($2 \leq n \leq N$)

$$ee_n = \frac{[L_n] - [R_n]}{[L_n] + [R_n]} \times 100. \quad (31)$$

A global or ensemble-averaged measure of the degree of symmetry breaking is provided by the number-weighted enantiomeric excess η :

$$\eta = \frac{\sum_{n=2}^N ([L_n] - [R_n])}{\sum_{n=2}^N ([L_n] + [R_n])} \times 100. \quad (32)$$

The importance of the enantiomeric excess is that it provides the order parameter for the symmetry breaking transition: the $|ee| \geq 0$ is zero for chiral symmetric states and nonzero otherwise. In the latter case, the Z_2 symmetry is broken. The average chain length of the homopolymers is given by:

$$\bar{n} = \frac{\sum_{n=2}^N n([L_n] + [R_n])}{\sum_{n=2}^N ([L_n] + [R_n])}, \quad (33)$$

and the root mean square deviation in the homochiral chain length are:

$$(\bar{n}^2)^{1/2} \equiv \sqrt{\langle (n - \bar{n})^2 \rangle} = \sqrt{\langle n^2 \rangle - \langle n \rangle^2}. \quad (34)$$

The angular brackets $\langle \rangle$ denote averaging with respect to the ensemble $\sum_n ([L_n] + [R_n])$, similar to eqn (33). It is important to remember that these are all time-dependent quantities.

Temporary but rather long lived asymmetric amplification can take place, as shown in Fig. 1; note the logarithmic time scale. The enantiomeric excess η averaged over all chain lengths eqn (32), from the dimer up to the maximum length chain $N = 12$ starts off at zero value until a time on the order of $t \sim 10$ at which the excess increases rapidly to nearly 100% at SMSB. This is followed by a gradual stepwise decrease or

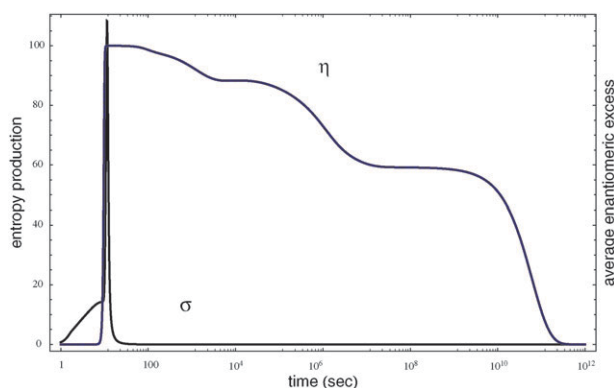


Fig. 1 Time evolution (logarithmic time scale) of the average enantiomeric excess η , averaged over all chains ($2 \leq n \leq 12$) eqn (32), and the associated entropy production σ , eqn (18). The entropy production peaks sharply at the onset of chiral symmetry breaking followed by a dramatic decrease to very small values. σ strictly approaches zero only at the racemization time scale $t_{\text{racem}} \gtrsim 10^{11}$. Initial concentrations: $[L_1]_0 = (1 \times 10^{-6} + 1 \times 10^{-15})$ M and $[R_1]_0 = 1 \times 10^{-6}$ M ($ee_0 = 5 \times 10^{-8}\%$), $[S]_0 = 2$ M, all other homo- and heterochiral oligomer initial concentrations are zero; and for the following rates $\varepsilon = 2 \times 10^{-5}$, $\varepsilon_- = 10^{-10}$, $k = 2.0$, $k_- = 10^{-5}$, $f = 0.9$, $k_{aa} = k_{bb} = 1.0$, $k_{ab} = k_{ba} = 10^3$, $k_{ab}^- = k_{ba}^- = 1.0$, $k_{aa}^- = k_{bb}^- = 10^{-5}$.

chiral erosion characterized by the appearance of quasi-plateaus of approximate constancy: η falls to about 90% at $t \sim 10^3$, then to about 60% at $t \sim 10^6$, staying approximately level until the final decrease to zero occurring at a time on the order of $t \sim 10^{11} - 10^{12}$. The system has racemized on this time scale. No appreciable differences in η can be discerned when we include the monomer, that is, start the sums at $n = 1$: we still observe slow chiral erosion proceeding through quasi-steady plateaus. The rate of the entropy produced eqn (18) exhibits an initial increase followed by a dramatic burst coinciding with the mirror symmetry breaking transition. This production then decreases rapidly to an exceedingly tiny but non zero value that remains constant during the entire period of slow chiral erosion, spanning more than ten orders of magnitude in time. The entropy production then goes strictly to zero when the system racemizes in complete accord with the fact that the system has reached chemical equilibrium. Although not displayed in Fig. 1, the substrate concentration falls from its initial value to zero approximately coincident with the peak structure of the entropy production, thus suggesting a connection between the sharp production of the latter and the change in S . The total entropy produced eqn (19) is $\sigma_T = 378.4 \text{ J mol}^{-1} \text{ K}^{-1}$. We can identify the major contributions to the entropy production in this process, see Fig. 2, from calculations of the partial entropies. In this way we find that the leading contribution comes from the monomer catalysis steps eqn (1), followed by the polymerization itself eqn (4), next by the mutual inhibition reactions: heterodimer formation eqn (6) and the “end-spoiled” cross-inhibition reactions eqn (5). The least important contribution comes from the direct production of monomers from the achiral substrate. The first three partial contributions all display a peak structure in the neighborhood of the symmetry breaking transition, with the corresponding peak values being displaced in time, see

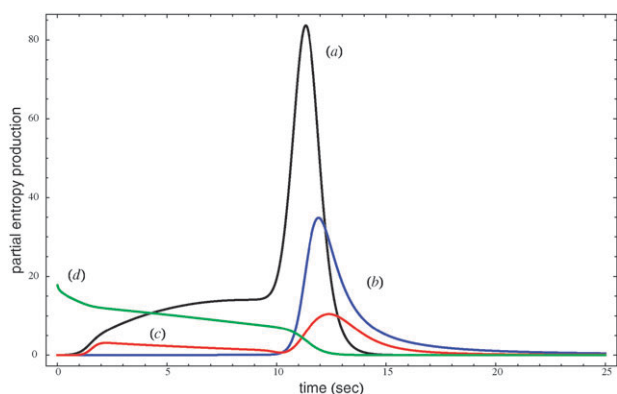


Fig. 2 Partial contributions to the total entropy production. From, (a) limited enantioselective monomer catalysis, (b) stepwise polymerization reactions, (c) chain-end termination and heterodimer formation reactions $\times 20$, and (d), direct monomer production $\times 1000$. Note how the peak values of (a), (b) and (c) are displaced in time. Same initial concentrations and rate constants as in Fig. 1.

Fig. 2. The exception to this is the entropy rate due to direct monomer production, which decreases monotonically.

A finer or more detailed measure of the degree of symmetry breaking and amplification is provided by the individual percent chain-length dependent enantiomeric excesses, eqn (31). A remarkable and complex dynamic behavior is revealed here. The time dependence of these n -dependent ee 's is plotted in Fig. 3; note the logarithmic time scale. The individual ee 's follow a common curve from initialization to chiral symmetry breaking, at about $t \sim 10$, and remain together at nearly 100% until about $t \sim 100$ at which time the common curve begins to split up into its constituents. Then, the percent chiral excess of each length homochiral chain behaves differently, until they again coalesce into a single curve upon final racemization, occurring at around $t \sim 10^{11}$. There is a common tendency for all the ee 's to decrease at intermediate time scales, with the largest length chains

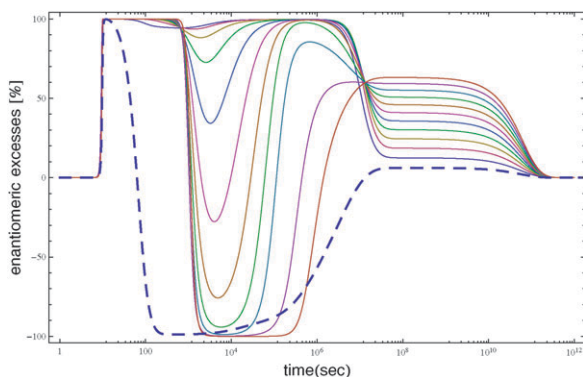


Fig. 3 Time dependence (logarithmic scale) of the individual chain-length dependent enantiomeric excesses $ee_n\% = \frac{[L_n] - [R_n]}{[L_n] + [R_n]} \times 100$, from the start of reactions to chiral symmetry breaking, and then on to the final racemization (family of solid curves). Near $t \simeq 10^4$, the sequence of curves from top to bottom corresponds to $n = 2$ to $n = 12$, respectively. Note damped oscillatory behavior of the excesses corresponding to $n = 7, 8, 9, 10, 11, 12$. The dashed curve shows the chiral excess for the monomers: $ee_1\%$. Same initial concentrations and rate constants as in Fig. 1.

(here, $N = 7, 8, 9, 10, 11, 12$) passing from *positive* then to *negative* values of the excess. The $N = 12, 11, 10$ chains exhibit nearly -100% excess during the period from 10^3 to 10^4 and beyond: there has been a chiral *sign reversal* in the excess corresponding to the largest chains. This holds also for the monomer ee_1 , plotted in the dashed curve. Except for the monomer, the individual excesses then all increase back to positive values at $t \sim 10^6$, then from $t \sim 10^7$ to $t \sim 10^{11}$, the excess increases sequentially as a function of the chain length n until racemization, where they all collapse to zero. The temporal behavior of the enantiomeric excesses of the largest chains $n = 12, 11, 10, 9, 8, 7$ is reminiscent of strongly *damped oscillations*. This oscillatory behavior in the enantiomeric excess can be understood in terms of the evolution of the individual concentrations of the longest chains. To illustrate this, we focus on the time dependence of the concentrations $[L_{12}]$ and $[R_{12}]$, the corresponding ee experiences the largest amplitude damped oscillations, see Fig. 4. For reference the inset diagram shows the enantiomeric excess over the entire time interval of the simulation, in comparison to Fig. 3. As shown in Fig. 4, the dominant concentration shifts back and forth between $[L_{12}]$ and $[R_{12}]$, respectively, until racemization when both concentrations converge to a common value. This leads to the chiral oscillations depicted in the inset. The shorter chains do not experience this oscillation, as illustrated for example by the concentrations $[L_5]$ and $[R_5]$ plotted in Fig. 5. There, the dominant concentration is always $[L_5]$, all the way from symmetry breaking at $t \sim 10$ to racemization, at approximately $t \gtrsim 10^{11}$. The corresponding ee suffers a dip near $t = 10^3$ (see inset), due to the concentration $[R_5]$ momentarily increasing at that time, see left hand graph of Fig. 5. This dip becomes more pronounced the longer the chain, see the sequence of curves around $t \sim 10^4$ in Fig. 3, and becomes a fully-fledged damped chiral oscillation for the longest chains in the system.

Static “snapshots” of this dynamic behavior nicely complement the evolution of the chain length dependent enantiomeric excesses. In Fig. 6 we display the enantiomeric excess *versus* the number of chiral repeat units at selected time slices. In the left-hand graph, the ee 's are all at 100% for all the chains. The next graph, corresponding to $t = 10^4$, shows the sign reversal tendency as a function of chain length, with the full reversal (-100%) being attained for the largest homochiral chains. The following graph, corresponding to $t = 10^9$ shows the monotonic increase of ee with chain length. Finally, the right-hand graph shows that racemization has set in by $t = 10^{13}$. These should be contrasted with Fig. 3. It is interesting to point out that both the qualitative behaviors depicted in the second and third snapshot have been reported in two recent and independent polymerization experiments. The tendency of the sign reversal in ee (from positive to negative values) as a function of chain length has been observed in the polymerization of racemic valine (Val-NCA) and leucine (Leu-NCA) in water subject to chiral initiators.²³ By contrast, the monotonic increase of the percent ee with chain length has been measured in independent chiral amplification experiments using leucine and glycine in water¹⁸ starting with a 20% initial enantiomeric excess of the L_1 monomer. These static snapshots also raise the important question of *when* to

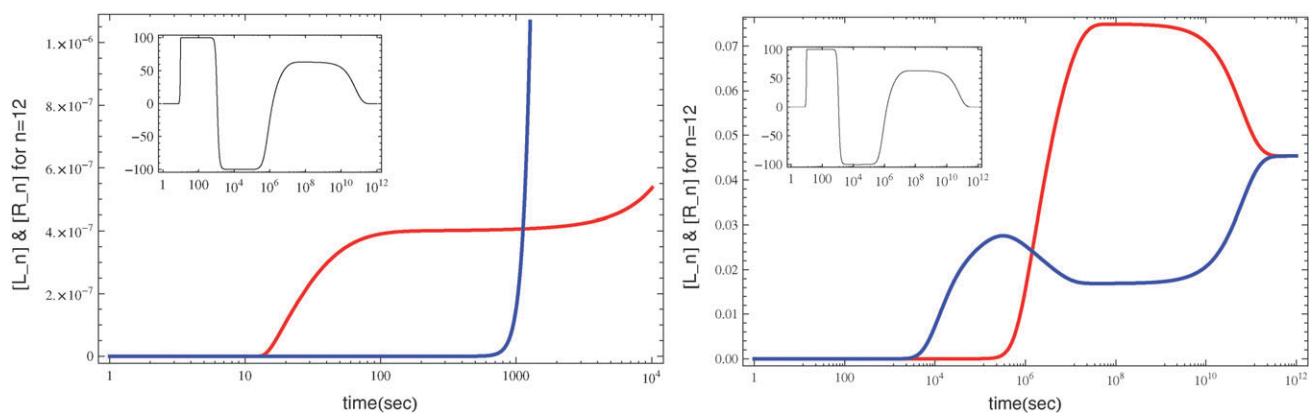


Fig. 4 Chiral oscillations. Evolution of the individual concentrations $[L_{12}]$ (the upper curve at the bifurcation) and $[R_{12}]$ for the complete time interval spanning symmetry breaking to final racemization. The inset graph gives the associated percent enantiomeric excess ee_{12} versus time. Left: expanded view of the initial stages of evolution for $1 \leq t \lesssim 10^4$. Symmetry breaking occurs at $t \sim 10$ with L_{12} dominating the excess. The chiral excess of these longest chains vanishes instantaneously for the *first* time at around $t = 10^3$ when the curves intersect, and then turns over such that now R_{12} dominates the chiral excess, see right hand graph (this leads to the sign flip in the excess, see inset). Right: chiral excess vanishes a *second* time at $t \sim 10^6$ when the two curves intersect again (compare to inset). Then from about 10^6 to 10^{11} the L_{12} chains again dominate the chiral excess until racemization, when the two curves collapse to a common curve (anti-bifurcation). Same initial concentrations and rate constants as in Fig. 1.

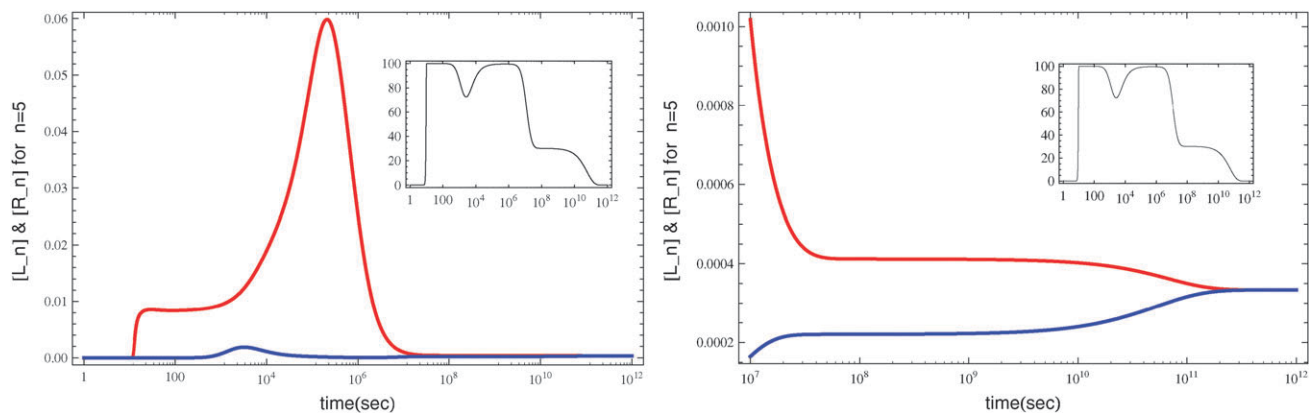


Fig. 5 Evolution of the individual concentrations $[L_5]$ (the upper curve at the bifurcation) and $[R_5]$ for the complete time interval spanning symmetry breaking to final racemization. The inset graph gives the associated percent enantiomeric excess ee_5 as a function of time. Left: full history of the evolution. Right: close up of the final time scales and racemization (anti-bifurcation).

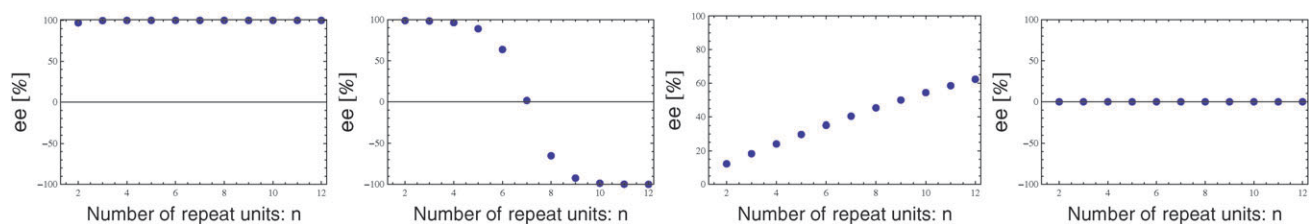


Fig. 6 Different time slices or “snap-shots” of Fig. 3 showing dependence of the chain-length dependent enantiomeric excesses $ee_n\% = \frac{[L_n] - [R_n]}{[L_n] + [R_n]} \times 100$ for $n \geq 2$, at different selected time scales. From left to right: total chiral symmetry breaking for all length homopolymer chains at $t = 100$, second graph shows the sign reversal tendency for the largest chains at $t = 10^4$, followed by the third graph, the monotone increase of chiral excess as a function of chain length at $t = 10^9$, and then the fourth graph, the final racemization at $t = 10^{13}$. Same initial concentrations and rate constants as in Fig. 1.

observe the chiral amplification and the enantiomeric excesses. In nonlinear reaction schemes such as this one, the enantiomeric excesses one measures can depend strongly on when the measurement or observation is made, that is, when one decides to *terminate* the experiment.

Additional information regarding the homo-oligomer composition of the chemical system is provided by the average homochiral chain length $\langle n \rangle$, see eqn (33). We plot this in Fig. 7 along with the standard deviation about the mean, eqn (34). The mean chain length starts off at 2,

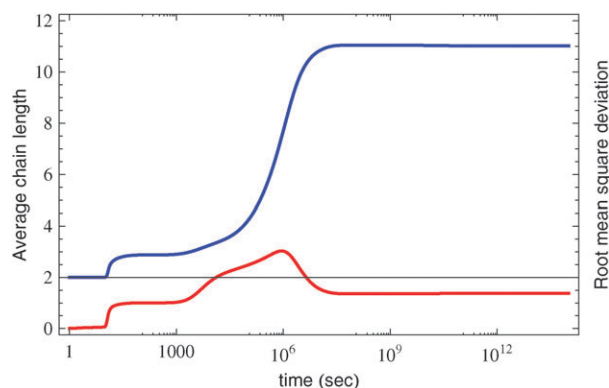


Fig. 7 The time evolution (logarithmic scale) of the average or mean homopolymer chain length \bar{n} (upper curve) and the corresponding root-mean-square deviation $(\overline{n^2})^{1/2}$ from the mean value (lower curve). The final stable values of the mean and RMS values are $\bar{n} = 11.02$ and $(\overline{n^2})^{1/2} = 1.38$, for $t \gtrsim 10^7$. This demonstrates that the final racemic mixture is dominated by the longer length homopolymer chains, and this is the final equilibrium configuration. Same initial concentrations and rate constants as in Fig. 1.

corresponding to the homodimer and then increases monotonically after the symmetry breaking transition, reaching a constant plateau at $t = 10^7$ where it remains constant all the way through to racemization and beyond. The final mean value $\langle n \rangle = 11.0$. The corresponding root mean square indicates that the fluctuations in the mean chain length increase as the mean chain length increases but then drops down to a constant value $(\overline{n^2})^{1/2} \sim 1.4$ when the average value stabilizes. This indicates that the final racemic composition is dominated by the longer chains: $n_{\text{final}} = 11.0 \pm 1.4$. The racemization time scale depends on how “irreversible” the model is. By way of example, if we increase the rate k_- of the reverse catalysis steps in eqn (1), keeping everything else constant, then the increased recycling of monomers back into achiral precursor S lowers this time scale as follows: $(k_-, t_{\text{racem}}) = (10^{-6}, 5 \times 10^{12}\text{s}), (10^{-5}, 5 \times 10^{11}\text{s}), (10^{-4}, 5 \times 10^{10}\text{s}), (10^{-3}, 1 \times 10^{10}\text{s}), (10^{-2}, 5 \times 10^9\text{s})$. By the same token, if we make k_- smaller, we can postpone racemization.

The enantiomeric cross inhibition $k_{ab} = k_{ba}$ is a determining factor in this model. By way of contrast, we consider a second $N = 12$ run with a much lower mutual inhibition than employed above, namely $k_{ab} = k_{ba} = 20$, and with the following inverse rates all set equal $k_{aa}^- = k_{bb}^- = k_{ab}^- = k_{ba}^- = 10^{-6}$, but keeping the remainder of the rates as before and with the same initial concentrations and excess. In this situation, the symmetry breaking occurs at a later time and most interestingly, the entropy production now peaks well *before* the mirror symmetry is broken, see Fig. 8. Fig. 9 shows that the catalysis still yields the major contribution to this peak, but the second and third most important contributions are now formation of end-chain spoiled oligomers followed by the polymerization, exactly opposite to the previous run employing the much higher mutual inhibition. The peak in σ is due principally to monomer catalysis, and not symmetry breaking.

The time dependence of these n -dependent ee 's is plotted in Fig. 10; note the logarithmic time scale. The individual ee 's

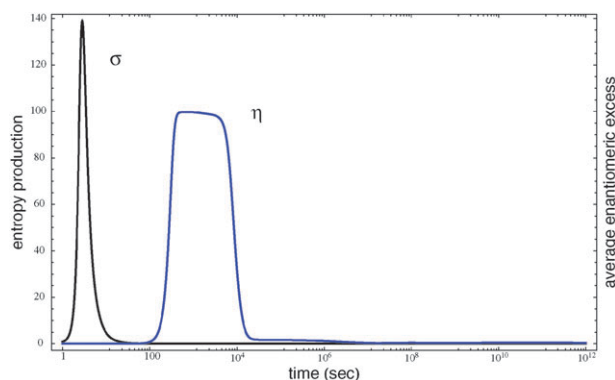


Fig. 8 Time evolution (logarithmic time scale) of the average enantiomeric excess η , averaged over all chains ($2 \leq n \leq 12$) eqn (32), and the associated entropy production σ , eqn (18). The entropy production peaks well before ($t \sim 300$) the onset of chiral symmetry breaking ($t \sim 300$) followed by a dramatic decrease to very small values. σ strictly approaches zero only at the racemization time scale $t_{\text{racem}} \gtrsim 10^{15}$. Initial concentrations: $[L_1]_0 = (1 \times 10^{-6} + 1 \times 10^{-15})$ M and $[R_1]_0 = 1 \times 10^{-6}$ M ($ee_0 = 5 \times 10^{-8}\%$), $[S]_0 = 2$ M, all other homo- and heterochiral oligomer initial concentrations are zero; and for the following rates $\varepsilon = 2 \times 10^{-5}$, $\varepsilon_- = 10^{-10}$, $k = 2.0$, $k_- = 10^{-5}$, $f = 0.9$, $k_{aa} = k_{bb} = 1.0$, $k_{ab} = k_{ba} = 20$, $k_{ab}^- = k_{ba}^- = k_{aa}^- = k_{bb}^- = 10^{-6}$.

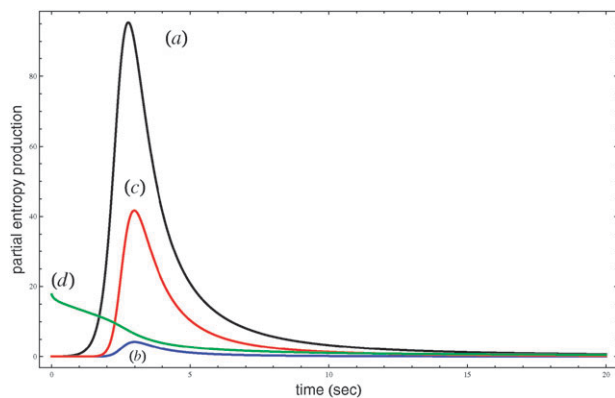


Fig. 9 Partial contributions to the total entropy production. From, (a) limited enantioselective monomer catalysis, (b) stepwise polymerization reactions, (c) chain-end termination and heterodimer formation reactions $\times 20$, and (d), direct monomer production $\times 1000$. Same initial concentrations and rate constants as in Fig. 8.

follow a common curve from initialization to chiral symmetry breaking, at about $t \sim 300$, and remain together at nearly 100% until about $t \sim 1000$ at which time the common curve begins to split up into its component parts. Note how the shorter homopolymers tend to racemize before the longer ones, there is a sequential chiral erosion in the individual enantiomeric excesses that is more pronounced the shorter the chain length. This holds as well for the monomer, plotted in the dashed curve (contrast to the monomer behavior in Fig. 3). Then, the percent excess of each length homochiral chain behaves differently, until they again coalesce into a single curve upon final racemization, occurring at around $t \sim 10^{14}$. The final approach to racemization is qualitatively very similar to the case treated above, compare the sequence in the right

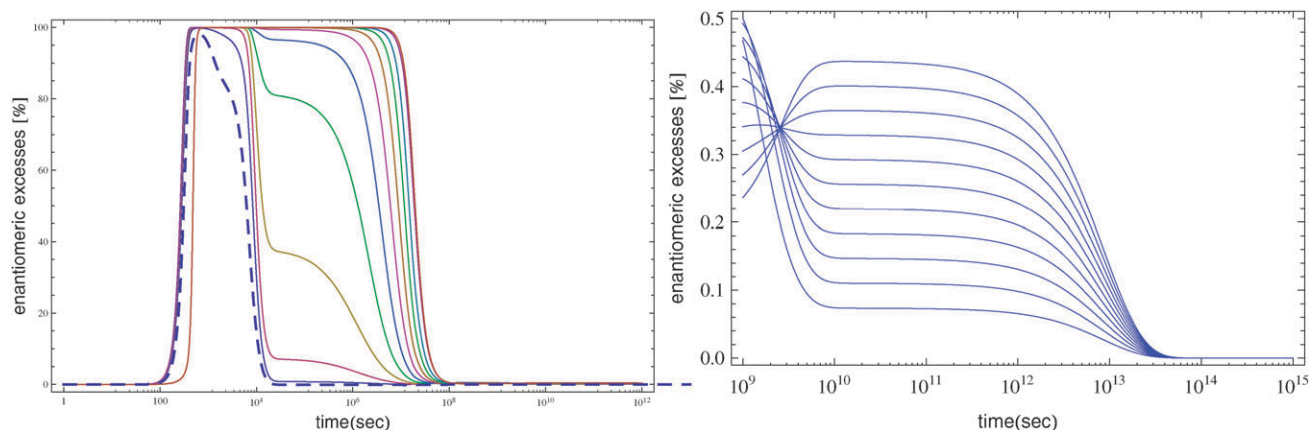


Fig. 10 Time dependence (logarithmic scale) of the individual chain-length dependent enantiomeric excesses $ee_n\% = \frac{[L_n] - [R_n]}{[L_n] + [R_n]} \times 100$, from the start of reactions to chiral symmetry breaking, and then on to the final racemization (family of solid curves). The dashed curve shows the chiral excess for the monomers. Right hand side shows a blow-up of the ee_n 's for the time scale $t = 10^9$ to 10^{15} , exhibiting the sequence of excesses and its final convergence to zero at racemization. Same initial concentrations and rate constants as in Fig. 8.

hand graph of Fig. 10; to the sequence of curves in Fig. 3 from roughly $t \sim 10^7$ to 10^{12} . A sequence of snapshots of the ee_n 's at selected times is displayed in Fig. 11.

Finally, we plot the average homochiral chain length $\langle n \rangle$, see eqn (33) in Fig. 12 along with the standard deviation. The mean chain length starts off at 2, corresponding to the homodimer and then increases monotonically after the symmetry breaking transition, reaching a constant plateau at about $t = 10^9$ where it remains constant all the way through to racemization and beyond. The final mean value $\langle n \rangle = 10.9$. Once again, the final racemic composition is dominated by the longest chains.

V. Conclusions and discussion

We have demonstrated that a strong chiral amplification can take place in a reversible model of chiral polymerization closed to matter flow and subject to constraints imposed by micro-reversibility. The inherent statistical fluctuations about the idealized racemic composition are modeled by an initial minuscule enantiomeric excess in a system dilute in the monomers. These results are important, because they suggest that spontaneous mirror symmetry breaking in experimental chiral polymerization can take place, and with observable and large chiral excesses without the need to introduce chiral initiators²⁴ or large initial chiral excesses.¹⁸ Instead, the

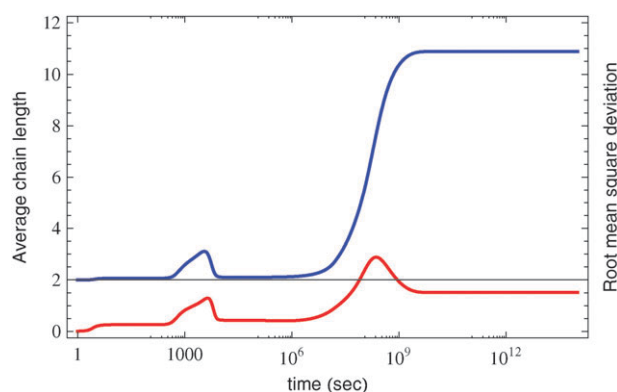


Fig. 12 The time evolution (logarithmic scale) of the average or mean homopolymer chain length \bar{n} (upper curve) and the corresponding root-mean-square deviation $\langle n^2 \rangle^{1/2}$ from the mean value (lower curve). The final stable values of the mean and RMS values are $\bar{n} = 10.9$ and $\langle n^2 \rangle^{1/2} = 1.5$, for $t \gtrsim 10^{10}$. The final racemic mixture is highly dominated by the longer length homopolymer chains, and this is the equilibrium configuration. Same initial concentrations and rate constants as in Fig. 8.

needed chiral monomers (*i.e.*, amino acids) can be produced directly from an achiral precursor and amplified *via* catalysis. Strong mutual inhibition is required to amplify the initial *ee* to large values, very similar to what we found for the reversible

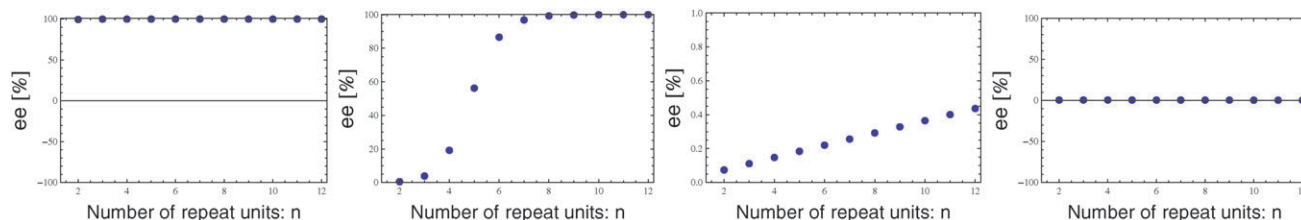


Fig. 11 Different time slices or “snap-shots” of Fig. 10 showing dependence of the chain-length dependent enantiomeric excesses $ee_n\% = \frac{[L_n] - [R_n]}{[L_n] + [R_n]} \times 100$ for $n \geq 2$, at different selected time scales. From left to right: total chiral symmetry breaking for all lengths of homopolymer chains at $t = 10^3$, the second graph shows the stepwise increase for the largest chains at $t = 10^6$, followed by the third graph, the monotonic increase of chiral excess as a function of chain length at $t = 10^{10}$, and then the fourth graph shows the final racemization at approximately $t = 10^{14}$. Same initial concentrations and rate constants as in Fig. 8.

Frank model in closed systems.⁴ The chain-length dependent enantiomeric excesses depend on time in a highly nontrivial way. The essential rate constant is that corresponding to the enantiomeric inhibition. A most intriguing novel feature revealed here for appreciable enantiomeric cross inhibition is the tendency for the chain length dependent enantiomeric excesses to exhibit a damped oscillatory behavior before the onset of final racemization. In these conditions, the observed chiral excess is clearly a time dependent phenomena, though the “period” of the chiral oscillations can be quite long. Oscillatory dynamics in chemical reactions has been observed experimentally, and analyzed theoretically and numerically in simple model systems;^{38,39} as far as we are aware, this behavior has not been revealed previously as a valid dynamical solution in polymerization models. The implications for chirality transmission are far reaching: “memory” of the sign of the initial chiral fluctuation is washed-out by the oscillations in the enantiomeric excess, adding another heretofore unexpected element of randomness to the process. While the sign of the initial chiral fluctuation is entirely random, any subsequent chiral oscillations can further “erase” the memory of the sign of this initial enantiomeric excess. These oscillations cease as the system approaches its equilibrium state. Moderate values lead to strong temporary symmetry breaking and larger values can lead to long period damped chiral oscillations before final racemization takes over.

We have also shown that the rate of entropy production per unit volume exhibits a peak value either before or near the vicinity of the chiral symmetry breaking transition. This increase to a peak value is mainly due to the catalytic production of the chiral monomers, followed next by the stepwise polymerization reactions, and then by the chain-end termination reactions and lastly, by direct monomer production. The rate falls to a vanishingly small but constant nonzero value maintained during the intermediate time scales, then drops to zero once the system has racemized. Previous calculations of the entropy produced in chiral symmetry breaking transitions have been carried out in the Frank model. In ref. 33, σ was evaluated for a reversible open flow Frank model with a constant inflow of achiral substrate and a constant outflow of the mutual inhibition product. In that situation, the entropy production rises from a small initial value and then levels off to a constant plateau after symmetry breaking. As the *open flow* keeps this system far from equilibrium, it can never racemize, and entropy is produced at a constant rate, *in sempiternum*. In ref. 34, the rate of entropy production was evaluated for a reversible open flow Frank model including limited enantioselectivity. Mirror symmetry is broken incompletely, and σ increases sharply from the start of the reaction, remains at a constant level during the lifetime of the initial racemic state and then decreases to a new stationary value once symmetry breaking sets in. These latter authors also consider the reversible Frank model with constant concentration of substrate but freely varying inhibition product. The entropy production increases, a peak value is reached when symmetry breaking is almost complete, then decreases to very small values. On the other hand, for a reversible Frank model *closed* to matter flow, and with strong mutual inhibition, the rate of entropy

production exhibits a sharp peak at the onset of symmetry breaking, falling to a tiny positive value until the system racemizes, at which time σ goes to zero.³⁷ This latter behavior is qualitatively similar to what we find in our polymerization model for large inhibition.

For the sake of computational simplicity, we have considered a model wherein no generally mixed heteropolymers were formed, only the heterodimers $LR = RL$ and LR_{n-1} and RL_{n-1} for $3 \leq n \leq N$. The corresponding number of differential equations grows linearly with the length N as $4N - 2$. A more realistic model should of course include the formation of all the possible heteropolymers of a given length n , *i.e.*, those heterochiral chains of length n containing $r \geq 1$ copies of L_1 and $s \geq 1$ copies of R_1 such that $r + s = n$. It is possible to build such a system, for example, starting with the copolymerization model of ref. 40, where the concentration variables are denoted as $c_{r,s}^L(t)|_{r \geq 1, s \geq 0}$ and $c_{r,s}^R(t)|_{r \geq 0, s \geq 1}$, the superscript indicating the final monomer in the chain, while the double subscript r, s encodes the number of individual L and R 's making up the chain, respectively. The number of differential equations in this case grows quadratically with polymer length N as $N(N+1)$, and is mathematically more involved. Detailed studies employing a *reversible* copolymerization scheme in closed systems will be presented elsewhere.⁴¹

Acknowledgements

We are grateful to Josep M. Ribó for a careful reading of the manuscript and for numerous useful discussions and to Michael Stich for his interest in oscillatory phenomena in chemical systems. This research is supported in part by the Grant AYA2009-13920-C02-01 from the Ministerio de Ciencia e Innovación (Spain) and forms part of the ESF COST Action CM07030: *Systems Chemistry*. CB is the recipient of a Calvo-Rodés training and research scholarship from the Instituto Nacional de Técnica Aeroespacial (INTA).

References

- 1 V. Avetisov and V. Goldanskii, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 11435.
- 2 A. Guijarro and M. Yus, *The Origin of Chirality in the Molecules of Life*, RSC Publishing, Cambridge, 2009.
- 3 I. Weissbuch, G. Bolbach, H. Zepik, E. Shavit, M. Tang, J. Frey, T. R. Jensen, K. Kajer, L. Leiserowitz and M. Lahav, *J. Am. Chem. Soc.*, 2002, **124**, 9093.
- 4 J. Crusats, D. Hochberg, A. Moyano and J. M. Ribó, *ChemPhysChem*, 2009, **10**(12), 2123.
- 5 D. K. Kondepudi, I. Progogine and G. Nelson, *Phys. Lett. A*, 1985, **111**, 29; A. Avetisov, V. V. Kuz'min and S. A. Aniki, *Chem. Phys.*, 1987, **112**, 179.
- 6 P. G. H. Sandars, *Origins Life Evol. Biosphere*, 2003, **33**, 575.
- 7 A. Brandenburg and T. Multamäki, *Int. J. Astrobiol.*, 2004, **3**, 209.
- 8 A. Brandenburg, A. C. Andersen, S. Höfner and M. Nilsson, *Origins Life Evol. Biosphere*, 2005, **35**, 225.
- 9 J. A. D. Wattis and P. V. Coveney, *Origins Life Evol. Biosphere*, 2005, **35**, 243.
- 10 M. Gleiser and J. Thorarinson, *Origins Life Evol. Biosphere*, 2006, **36**, 501.
- 11 M. Gleiser, *Origins Life Evol. Biosphere*, 2007, **37**, 235.
- 12 M. Gleiser and S. I. Walker, *Origins Life Evol. Biosphere*, 2008, **38**, 293.

-
- 13 M. Gleiser, J. Thorarinson and S. I. Walker, *Origins Life Evol. Biosphere*, 2008, **38**, 499.
 - 14 Y. Saito and J. Hyuga, *J. Phys. Soc. Jpn.*, 2005, **74**, 1629.
 - 15 F. C. Frank, *Biochim. Biophys. Acta*, 1953, **11**, 459.
 - 16 R. Plasson, H. Bersini and A. Commeyras, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 16733.
 - 17 M. Blocher, T. Hitz and P. L. Luisi, *Helv. Chim. Acta*, 2001, **84**, 375.
 - 18 T. Hitz and P. L. Luisi, *Helv. Chim. Acta*, 2003, **86**, 1423.
 - 19 J. G. Nery, G. Bolbach, I. Weissbuch and M. Lahav, *Angew. Chem., Int. Ed.*, 2003, **42**, 2157.
 - 20 T. Hitz and P. L. Luisi, *Origins Life Evol. Biosphere*, 2004, **34**, 93.
 - 21 I. Rubinstein, R. Eliash, G. Bolbach, I. Weissbuch and M. Lahav, *Angew. Chem., Int. Ed.*, 2007, **46**, 3710.
 - 22 R. A. Illos, F. R. Bisogno, G. Clodic, G. Bolbach, I. Weissbuch and M. Lahav, *J. Am. Chem. Soc.*, 2008, **130**, 8651.
 - 23 I. Rubinstein, G. Clodic, G. Bolbach, I. Weissbuch and M. Lahav, *Chem.–Eur. J.*, 2008, **14**, 10999.
 - 24 I. Weissbuch, R. A. Illos, G. Bolbach and M. Lahav, *Acc. Chem. Res.*, 2009, **42**(8), 1128.
 - 25 K. Mislow, *Collect. Czech. Chem. Commun.*, 2003, **68**, 849.
 - 26 J. S. Siegel, *Chirality*, 1998, **10**, 24.
 - 27 U. Stier, *J. Chem. Phys.*, 2006, **124**, 024901; U. Stier, *J. Chem. Phys.*, 2006, **125**, 024903.
 - 28 D. Hochberg, *Chem. Phys. Lett.*, 2010, **491**, 237.
 - 29 A. Brandenburg, A. C. Andersen and M. Nilsson, *Origins Life Evol. Biosphere*, 2005, **35**, 507.
 - 30 P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics*, Cambridge University Press, Cambridge, 1995, ch. 3.
 - 31 T. F. A. De Greef, M. M. J. Smulders, M. Wolffs, A. P. H. J. Schenning, R. P. Sijbesma and E. W. Meijer, *Chem. Rev.*, 2009, **109**(11), 5687.
 - 32 J. Ross and M. O. Vlad, *J. Phys. Chem. A*, 2005, **109**, 10607.
 - 33 D. Kondepudi and L. Kapcha, *Chirality*, 2008, **20**, 524.
 - 34 M. Mauksch and S. B. Tsogoeva, *ChemPhysChem*, 2008, **9**, 2359.
 - 35 D. Kondepudi and I. Prigogine, *Modern Thermodynamics: From Heat engines to dissipative structures*, New York, Wiley, 1998.
 - 36 W. H. Mills, *J. Soc. Chem. Ind.*, 1932, **10**, 750; V. I. Goldanski and V. V. Kuzmin, *Z. Phys. Chem. (Leipzig)*, 1988, **269**, 269.
 - 37 J. M. Ribó, private communication.
 - 38 S. K. Scott, *Oscillations, Waves, and Chaos in Chemical Kinetics*, Oxford University Press, Oxford, 1994.
 - 39 M. C. Cross and P. C. Hohenberg, *Rev. Mod. Phys.*, 1993, **65**, 851.
 - 40 J. A. Wattis and P. V. Coveney, *J. Phys. Chem. B*, 2007, **111**, 9546.
 - 41 C. Blanco and D. Hochberg, work in progress.

Cite this: *Phys. Chem. Chem. Phys.*, 2012, **14**, 2301–2311

www.rsc.org/pccp

PAPER

Homochiral oligopeptides by chiral amplification: interpretation of experimental data with a copolymerization model

Celia Blanco and David Hochberg*

Received 5th September 2011, Accepted 7th December 2011

DOI: 10.1039/c2cp22813k

We present a differential rate equation model of chiral polymerization based on a simple copolymerization scheme in which the enantiomers are added to, or removed from, the homochiral or heterochiral chains (reversible stepwise isodesmic growth or dissociation). The model is set up for closed systems and takes into account the corresponding thermodynamic constraints implied by the reversible monomer attachments, while obeying a constant mass constraint. In its simplest form, the model depends on a single variable rate constant, the maximum chain length N , and the initial concentrations. We have fit the model to the experimental data from the Rehovot group on lattice-controlled chiral amplification of oligopeptides. We find in all the chemical systems employed, except for one, that the model fits the measured relative abundances of the oligopeptides with higher degrees of correlation than from a purely random polymerization process.

1 Introduction

In the transition from prebiotic racemic chemistry to chiral biology one scenario suggests that homochiral peptides must have appeared before the appearance of the primeval enzymes.^{1,2} While several stochastic synthetic routes for mirror symmetry breaking that convert racemates into nonracemates have been described,^{3,4} the generation of long bio-like polymers¹ made up of repeating units of the same handedness requires elaboration of new synthetic routes. Polymerization reactions of racemic mixtures of monomers in solution are typically expected to yield polymers composed of random sequences of the left- and right-handed repeat units following a binomial or Bernoulli distribution. Thus the probability of obtaining oligomers with a homochiral sequence becomes negligible with increasing length.¹

Recent investigations have proposed that N -carboxyanhydride (NCA)^{5,6} and thioester derivatives^{7,8} of amino acids might have operated as relevant precursors in the formation of the early peptides.⁹ Results on the polymerization of NCA monomers in organic solvents,^{10–15} in water^{16–18} and in the solid state^{19,20} have been published. Luisi and coworkers^{21–24} have reported the polymerization of racemic α -amino acids in solution which yields small amounts of oligopeptides of homochiral sequence whose abundances with respect to the heterochiral chains exhibit a slight departure from the binomial distribution.

This problem of the random distribution can be overcome by catalyzed polymerization of amphiphilic amino acids, in

racemic and nonracemic forms, which self-assemble into two-dimensional ordered crystallites at the air–water interface.^{25,26} Based on a process involving self-assembly followed by lattice controlled polymerization, Lahav and coworkers recently proposed a general scenario for the generation of homochiral oligopeptides of a single handedness from non-racemic mixtures of activated alpha amino acids.^{25,26} Initial non-racemic mixtures undergo a phase separation by self-assembly into a 2D racemic crystalline phase and a separate enantiomorphous 2D phase of the enantiomer in excess. Each of these crystalline phases has markedly different chemical properties, thus yielding products that differ in the composition of the oligomers. So, polymerization within the enantiomorphous crystalline phase yields homochiral oligopeptides of one handedness whereas the reaction controlled by the racemic crystallites yields racemic mixtures and heterochiral products. The combination of the two routes leads to an overall chiral amplification process.

In this paper, we are interested in the lattice-controlled polymerization reactions proposed by those authors. It is important to clarify at the outset what specific aspect of the overall experimental mechanism we want to model here and the way we aim to do so. The proposed experimental scheme starts from an initial excess, say $S > R$ of monomers which undergoes an initial self-assembly process into two types of two-dimensional crystallites at the air/water interface. Once formed, each one of these two crystal phases participates in the control of a subsequent type of polymerization. Thus, the racemic crystallites polymerize racemic mixtures of oligomers and the heterochiral products, whereas the other pure enantiomorphous crystallite controls the polymerization of the isotactic chains, these are formed from the monomer in excess

Centro de Astrobiología (CSIC-INTA), Ctra. Ajalvir Km. 4, 28850 Torrejón de Ardoz, Madrid, Spain.
E-mail: blancodc@cab.inta-csic.es, hochbergd@cab.inta-csic.es

(*S*, in this example). However, the details of the polymerizations depend in a complicated way upon the specific *packing arrangements* of the crystal monomers and the possible *reaction pathways* taken within each crystallite phase. The authors of the experiments state that the connection between the monomer packing arrangements in the crystallites and the resultant composition of the various diastereoisomeric products is “not straightforward”.²⁶ We therefore opt for a simple model for interpreting their data. With this objective in mind, we present a copolymerization model for the interpretation of the experimental data. The model may be termed *effective* in the following sense: it presupposes or takes as given the *prior* formation of the self-assembled 2D crystallites at the air–water interface and is concerned exclusively with the subsequent polymerization reactions. Thus the complicated microscopic details referring to the monomer packing arrangements and reaction pathways within the crystallite self-assemblies are treated implicitly with our rate constants. Our copolymerization reaction rates can satisfactorily account for the different chemical properties of the two crystalline phases (racemic 2D crystallites and pure enantiomorphous 2D crystallites) that lead to the formation of racemic mixtures, heterochiral products and isotactic oligopeptides. We contrast the fits from our model with those assuming a purely random process that obeys a binomial distribution. The final justification for considering such an effective model rests on its ability to yield good fits to the data. The goodness of the fits obtained below demonstrates that the experimental data can be fit convincingly as if the simple scheme depicted pictorially in Fig. 1 were the sole mechanism leading to the observed relative abundances. This then gives additional meaning to the term “effective”, and in the operational sense.

2 The copolymerization model

Our starting point is a simple model for the copolymerization of two chemically distinct monomers displaying a wide variety of product sequence compositions. The model we introduce and study here is an appropriately modified and extended version of the one considered a few years ago by Wattis and Coveney.²⁷

The main important differences compared to prior and related models are that (1) we consider polymerization in *closed* systems²⁸—so that no matter flow is permitted with an external environment—and (2) we allow for reversible monomer association steps. We also correctly include the formation (and dissociation) of the heterodimer.²⁸ It turns out that this must be treated on a separate basis in order to avoid double counting, which if left unchecked, would lead to a violation in the constant mass constraint. Once the heterodimer is treated correctly, this implies that the hetero-trimer must also be treated separately. Besides this, the remainder of the hetero-oligomers can be treated in a uniform way.

First, we introduce the notation to be used. Polymers are classified by three quantities: the number of *A* monomers of which it is composed (subscript *r*), the number of *B* monomers which it contains (subscript *s*) and the final or terminal monomer in the chain, denoted by a superscript. In this scheme, the monomers are denoted by $A = C_{1,0}^A$ and $B = C_{0,1}^B$; pure homopolymers are denoted by $C_{r,0}^A$ and $C_{0,s}^B$; all copolymer chains $C_{r,s}^A$ or $C_{r,s}^B$ with $r, s \geq 1$ are heteropolymers. Note also that chains of the form $C_{0,s}^A$ and $C_{r,0}^B$

• Homopolymerization



• Heterodimerization



• Heterotrimerization

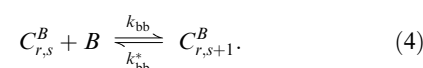
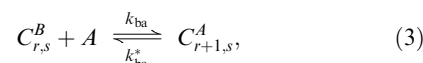
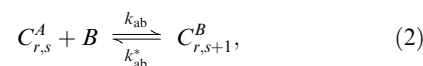
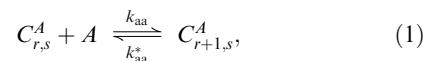


• Heteropolymerization



Fig. 1 The copolymerization model. The (R)-chiral (red) and (S)-chiral (blue) monomers reversibly associate into the growing homochiral (top) or heterochiral (bottom) copolymer chains. Because the system is *closed*, both the heterodimer (second line) and heterotrimer (third and fourth lines) reactions must be treated separately to avoid double counting and thus ensure that the total system mass is conserved in a closed system (see the text for an explanation).

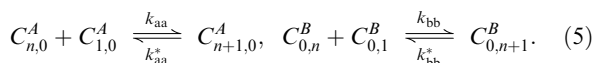
are forbidden. The corresponding time-dependent concentrations are denoted by lower case variables: e.g., $c_{r,s}^A(t)$ and $c_{r,s}^B(t)$. The model is then defined by the following reactions, in which equilibrium is maintained between the finite monomer pool and the ensemble of copolymers:



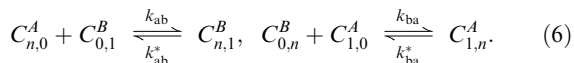
This model can accommodate any two chemically distinct monomers. For the purpose of this paper, we consider the case when $A = R$ and $B = S$ are two enantiomers.

The overall basic scheme must be broken down into several special subcases, especially important so as to avoid undesired double counting of the heterodimer and heterotrimer reactions, see Fig. 1. Once we treat these special cases, we then pass to the corresponding set of rate equations for the concentrations.

The formation of chirally pure polymer chains denoted by $C_{n,0}^A$ and $C_{0,n}^B$ for $1 \leq n \leq N - 1$ is described by the homopolymerization reactions:

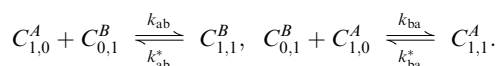


where N is the maximum chain length permitted. In our recently reported work,²⁸ we considered that once a monomer has been added to a homopolymer of the opposite chirality (that is, “the wrong” monomer), the polymer is inhibited and further growth is halted. This polymer could not directly react anymore and could only lose its wrong terminal monomer through the inverse reaction. In the present model, we assume that such a chain can continue to grow by adding monomers of both configurations. So, for $2 \leq n \leq N - 1$, the heteropolymerization or inhibition reactions are as follows:

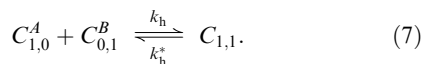


For both homo- and hetero-polymerization reactions, represented by eqn (5) and (6), the upper limits specified for n ensure that the *maximum* length for all oligomers produced (or consumed) by these reaction sets, both the homo- and heterochiral ones, is never greater than N . In the remainder of this paper we will consider here the natural and chiral symmetric reaction rate assignments $k_{aa} = k_{bb}$, $k_{ab} = k_{ba}$ and likewise for the inverse rates, $k_{aa}^* = k_{bb}^*$ and $k_{ab}^* = k_{ba}^*$, reducing the number of independent rate constants to four.

Even if we have information about the composition, we can only know the chirality of the last monomer attached to the chain, we have no information regarding the specific *sequence*. This implies that the following two reactions are indistinguishable:



Thus for all practical purposes, $C_{1,1}^A \equiv C_{1,1}^B$, and this suggests using the following notation: $C_{1,1} \equiv C_{1,1}^A \equiv C_{1,1}^B$, and to define a unique direct constant rate: $k_h = \frac{k_{ab} + k_{ba}}{2}$, and an inverse one $k_h^* = \frac{k_{ab}^* + k_{ba}^*}{2}$. Note that if $k_{ab} = k_{ba}$, then $k_h = k_{ab} = k_{ba}$. Due to these characteristics, we will treat the heterodimer in a different way compared with the other hetero-polymers. The reaction of the heterodimer formation is therefore:

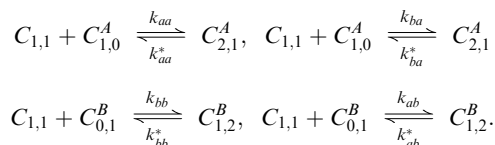


As before, the reactants and products in eqn (7) are the same, so the differences in the free energy between initial and final states should be the same in all the reactions in these

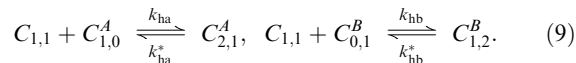
equations, implying the following thermodynamic constraint on the reaction rates:

$$\frac{k_{ab}}{k_{ab}^*} = \frac{k_{ba}}{k_{ba}^*}. \quad (8)$$

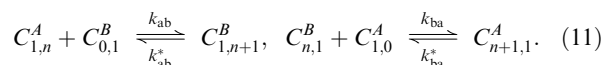
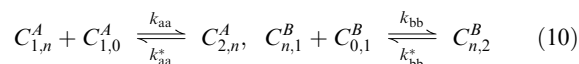
If the heterodimer formation were not to be treated in a separate way as we have done, and were to be included, *e.g.* in eqn (6) by merely changing the lower limits for n ($2 \leq n \leq N - 1$) by $1 \leq n \leq N - 1$, we would be making the mistake of double counting it. The same occurs for the heteropolymers formed from the addition of a monomer to a heterodimer. The two reactions of each pair of the following equations are also indistinguishable:



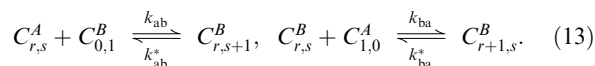
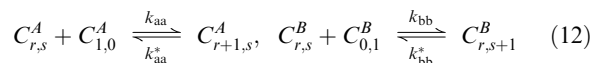
Again, it is convenient to define the following direct reaction rates for these steps, $k_{ha} = \frac{k_{aa} + k_{ba}}{2}$, $k_{hb} = \frac{k_{bb} + k_{ab}}{2}$ and inverse rates: $k_{ha}^* = \frac{k_{aa}^* + k_{ba}^*}{2}$, $k_{hb}^* = \frac{k_{bb}^* + k_{ab}^*}{2}$. Note that if $k_{aa} = k_{bb}$ and $k_{ab} = k_{ba}$, then $k_{ha} = k_{hb}$, and if $k_{aa}^* = k_{bb}^*$ and $k_{ab}^* = k_{ba}^*$, then $k_{ha}^* = k_{hb}^*$. The reactions to be considered are then:



As we have already remarked, in our model, as in the original one for open systems,²⁷ the polymeric chains that have taken up the “wrong” chirality monomer can continue to grow. Thus, we allow for the further growth of these chains by adding monomers of either chirality. This kind of polymerization reaction for $2 \leq n \leq N - 2$ is given by:



And for $2 \leq r \leq N - 2$, $1 \leq s \leq N - 1 - r$:



Note that in the elementary reaction steps, in the rate constants, and in the corresponding differential rate equations (see below), the left–right symmetry of the model is manifested, that is, it possesses a discrete Z_2 symmetry. This symmetry can be broken spontaneously by the dynamical solutions of the differential rate equations, thus this model is apt for studying spontaneous mirror symmetry breaking.

By lifting the Z_2 degeneracy in the reaction rates, *e.g.*, allowing for $k_{aa} \neq k_{bb}$ and thus leading to more independent rate constants for describing the reaction set, we could study the influence of explicit chiral bias on the model. As this is not the aim of this work, we will not consider it here.

We next write down the differential rate equations corresponding to this reaction network. We employ the rate-equation theory as in chemical kinetics. We begin with the rate equations for the chiral monomers:

$$\begin{aligned} \frac{dc_{1,0}^A}{dt} = & -k_{aa}c_{1,0}^A \left(2c_{1,0}^A + \sum_{n=2}^{N-1} c_{n,0}^A + \sum_{n=2}^{N-2} c_{1,n}^A + \sum_{r=2}^{N-2} \sum_{s=1}^{N-1-r} c_{r,s}^A \right) \\ & - k_{ba}c_{1,0}^A \left(\sum_{n=2}^{N-1} c_{0,n}^B + \sum_{n=2}^{N-2} c_{n,1}^B + \sum_{s=2}^{N-2} \sum_{r=1}^{N-1-s} c_{r,s}^B \right) \\ & - k_h c_{1,0}^A c_{0,1}^B - k_{ha} c_{1,0}^A c_{1,1} \\ & + k_{aa}^* \left(2c_{2,0}^A + \sum_{n=3}^N c_{n,0}^A + \sum_{n=2}^{N-2} c_{2,n}^A + \sum_{r=3}^{N-1} \sum_{s=1}^{N-r} c_{r,s}^A \right) \\ & + k_{ba}^* \left(\sum_{n=2}^{N-1} c_{1,n}^A + \sum_{n=2}^{N-2} c_{2,n}^A + \sum_{r=3}^{N-1} \sum_{s=1}^{N-r} c_{r,s}^A \right) \\ & + k_h^* c_{1,1} + k_{ha}^* c_{2,1} \end{aligned} \quad (14)$$

$$\begin{aligned} \frac{dc_{0,1}^B}{dt} = & -k_{bb}c_{0,1}^B \left(2c_{0,1}^B + \sum_{n=2}^{N-1} c_{0,n}^B + \sum_{n=2}^{N-2} c_{n,1}^B + \sum_{s=2}^{N-2} \sum_{r=1}^{N-1-s} c_{r,s}^B \right) \\ & - k_{ab}c_{0,1}^B \left(\sum_{n=2}^{N-1} c_{n,0}^A + \sum_{n=2}^{N-2} c_{1,n}^A + \sum_{r=2}^{N-2} \sum_{s=1}^{N-1-r} c_{r,s}^A \right) \\ & - k_h c_{1,0}^A c_{0,1}^B - k_{hb} c_{0,1}^B c_{1,1} \\ & + k_{bb}^* \left(2c_{0,2}^B + \sum_{n=3}^N c_{0,n}^B + \sum_{n=2}^{N-2} c_{n,2}^B + \sum_{s=3}^{N-1} \sum_{r=1}^{N-s} c_{r,s}^B \right) \\ & + k_{ab}^* \left(\sum_{n=2}^{N-1} c_{n,1}^B + \sum_{n=2}^{N-2} c_{n,2}^B + \sum_{s=3}^{N-1} \sum_{r=1}^{N-s} c_{r,s}^B \right) \\ & + k_h^* c_{1,1} + k_{hb}^* c_{1,2} \end{aligned} \quad (15)$$

The equations describing the concentration of the homopolymers, for $2 \leq n \leq N-1$, are:

$$\begin{aligned} \frac{dc_{n,0}^A}{dt} = & k_{aa}c_{1,0}^A (c_{n-1,0}^A - c_{n,0}^A) - k_{ab}c_{n,0}^A c_{0,1}^B \\ & + k_{aa}^* (c_{n+1,0}^A - c_{n,0}^A) + k_{ab}^* c_{n,1}^B \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{dc_{0,n}^B}{dt} = & k_{bb}c_{0,1}^B (c_{0,n-1}^B - c_{0,n}^B) - k_{ba}c_{0,n}^B c_{1,0}^A \\ & + k_{bb}^* (c_{0,n+1}^B - c_{0,n}^B) + k_{ba}^* c_{1,n}^A \end{aligned} \quad (17)$$

It is necessary to treat the kinetic equations of the maximum length homopolymers N individually. Since these do not elongate further, they cannot directly react, and cannot be the product of an inverse reaction involving a longer chain:

$$\frac{dc_{N,0}^A}{dt} = k_{aa}c_{1,0}^A c_{N-1,0}^A - k_{aa}^* c_{N,0}^A \quad (18)$$

$$\frac{dc_{0,N}^B}{dt} = k_{bb}c_{0,1}^B c_{0,N-1}^B - k_{bb}^* c_{0,N}^B \quad (19)$$

The differential equations describing the concentration of each type of heteropolymer (included the heterodimer), for $2 \leq n \leq N-2$, are:

$$\begin{aligned} \frac{dc_{1,1}^A}{dt} = & k_h c_{1,0}^A c_{0,1}^B - k_{ha} c_{1,1} c_{1,0}^A - k_{hb} c_{1,1} c_{0,1}^B \\ & - k_h^* c_{1,1} + k_{ha}^* c_{2,1}^A + k_{hb}^* c_{1,2}^B \end{aligned} \quad (20)$$

$$\begin{aligned} \frac{dc_{1,n}^A}{dt} = & -k_{aa}c_{1,0}^A c_{1,n}^A - k_{ab}c_{0,1}^B c_{1,n}^A + k_{ba}c_{0,n}^B c_{1,0}^A \\ & + k_{aa}^* c_{2,n}^A + k_{ab}^* c_{1,n+1}^B - k_{ba}^* c_{1,n}^A \end{aligned} \quad (21)$$

$$\begin{aligned} \frac{dc_{n,1}^B}{dt} = & -k_{bb}c_{0,1}^B c_{n,1}^B - k_{ba}c_{1,0}^A c_{n,1}^B + k_{ab}c_{n,0}^A c_{0,1}^B \\ & + k_{bb}^* c_{n,2}^B + k_{ba}^* c_{n+1,1}^A - k_{ab}^* c_{n,1}^B \end{aligned} \quad (22)$$

As before, it is useful to treat individually the maximum length polymers N :

$$\frac{dc_{1,N-1}^A}{dt} = k_{ba}c_{0,N-1}^B c_{1,0}^A - k_{ba}^* c_{1,N-1}^A \quad (23)$$

$$\frac{dc_{N-1,1}^B}{dt} = k_{ab}c_{N-1,0}^A c_{0,1}^B - k_{ab}^* c_{N-1,1}^B \quad (24)$$

As was mentioned when describing the reaction network, each kind of trimer $c_{2,1}^A$ and $c_{1,2}^B$ must have its own differential equation in terms of k_{ha} , k_{hb} :

$$\begin{aligned} \frac{dc_{2,1}^A}{dt} = & -k_{aa}c_{1,0}^A c_{2,1}^A - k_{ab}c_{0,1}^B c_{2,1}^A + k_{ha}c_{1,1} c_{1,0}^A \\ & + k_{aa}^* c_{3,1}^A + k_{ab}^* c_{2,2}^B - k_{ha}^* c_{2,1}^A \end{aligned} \quad (25)$$

$$\begin{aligned} \frac{dc_{1,2}^B}{dt} = & -k_{bb}c_{0,1}^B c_{1,2}^B - k_{ba}c_{1,0}^A c_{1,2}^B + k_{hb}c_{1,1} c_{0,1}^B \\ & + k_{bb}^* c_{1,3}^B + k_{ba}^* c_{2,2}^A - k_{hb}^* c_{1,2}^B \end{aligned} \quad (26)$$

For $2 \leq n \leq N-3$:

$$\begin{aligned} \frac{dc_{2,n}^A}{dt} = & k_{aa}c_{1,0}^A (c_{1,n}^A - c_{2,n}^A) - k_{ab}c_{0,1}^B c_{2,n}^A + k_{ba}c_{1,n}^B c_{1,0}^A \\ & + k_{aa}^* (c_{3,n}^A - c_{2,n}^A) + k_{ab}^* c_{2,n+1}^B - k_{ba}^* c_{2,n}^A \end{aligned} \quad (27)$$

$$\begin{aligned} \frac{dc_{n,2}^B}{dt} = & k_{bb}c_{0,1}^B (c_{n,1}^B - c_{n,2}^B) - k_{ba}c_{1,0}^A c_{n,2}^B + k_{ab}c_{n,1}^A c_{0,1}^B \\ & + k_{bb}^* (c_{n,3}^B - c_{n,2}^B) + k_{ba}^* c_{n+1,2}^A - k_{ab}^* c_{n,2}^B \end{aligned} \quad (28)$$

Once again, the equations corresponding to the maximum length homopolymers N are:

$$\begin{aligned} \frac{dc_{2,N-2}^A}{dt} = & k_{aa}c_{1,0}^A c_{1,N-2}^A + k_{ba}c_{1,N-2}^B c_{1,0}^A \\ & - k_{aa}^* c_{2,N-2}^A - k_{ba}^* c_{2,N-2}^A \end{aligned} \quad (29)$$

$$\begin{aligned} \frac{dc_{N-2,2}^B}{dt} = & k_{bb}c_{0,1}^B c_{N-2,1}^B + k_{ab}c_{N-2,1}^A c_{0,1}^B \\ & - k_{bb}^* c_{N-2,2}^B - k_{ab}^* c_{N-2,2}^B \end{aligned} \quad (30)$$

Table 1 Number of differential equations as a function of the maximum polymer length N

	Number of eqn		Number of eqn
$c_{1,0}^A$	1	$c_{0,1}^B$	1
$c_{1,1}^A$	1	$c_{0,n}^B, (2 \leq n \leq N)$	$\sum_{n=2}^N = N - 1$
$c_{n,0}^A, (2 \leq n \leq N)$	$\sum_{n=2}^N = N - 1$	$c_{n,1}^B, (2 \leq n \leq N - 1)$	$\sum_{n=2}^{N-1} = N - 2$
$c_{1,n}^A, (2 \leq n \leq N - 1)$	$\sum_{n=2}^{N-1} = N - 2$	$c_{1,2}^B$	1
$c_{2,1}^A$	1	$c_{n,2}^B, (2 \leq n \leq N - 2)$	$\sum_{n=2}^{N-2} = N - 3$
$c_{2,n}^A, (2 \leq n \leq N - 2)$	$\sum_{n=2}^{N-2} = N - 3$	$c_{r,s}^B, (3 \leq s \leq N - 2)$	$\sum_{r=1}^{N-2} \sum_{s=3}^{N-2} = \frac{1}{2}(N^2 - 7N + 12)$
$c_{r,s}^A, (3 \leq r \leq N - 2)$	$\sum_{r=3}^{N-2} \sum_{s=1}^{N-1-r} = \frac{1}{2}(N^2 - 7N + 12)$	$c_{r,s}^B, (1 \leq r \leq N - 1 - s)$	$\sum_{r=1}^{N-1-s} \sum_{s=3}^{N-2} = \frac{1}{2}(N^2 - 7N + 12)$
$(1 \leq s \leq N - 1 - r)$		$c_{n,N-n}^B, (3 \leq n \leq N - 1)$	$\sum_{n=3}^{N-1} = N - 3$
$c_{n,N-n}^A, (3 \leq n \leq N - 1)$	$\sum_{n=3}^{N-1} = N - 3$		

For $3 \leq r \leq N - 2$ and $1 \leq s \leq N - 1 - r$:

$$\begin{aligned} \frac{dc_{r,s}^A}{dt} = & k_{aa}c_{1,0}^A(c_{r-1,s}^A - c_{r,s}^A) - k_{ab}c_{0,1}^Bc_{r,s}^A + k_{ba}c_{r-1,s}^Ac_{1,0}^A \\ & + k_{aa}^*(c_{r+1,s}^A - c_{r,s}^A) + k_{ab}^*c_{r,s+1}^B - k_{ba}^*c_{r,s}^A \end{aligned} \quad (31)$$

For $3 \leq s \leq N - 2$ and $1 \leq r \leq N - 1 - s$:

$$\begin{aligned} \frac{dc_{r,s}^B}{dt} = & k_{bb}c_{0,1}^B(c_{r,s-1}^B - c_{r,s}^B) - k_{ba}c_{1,0}^Ac_{r,s}^B + k_{ab}c_{r,s-1}^Ac_{0,1}^B \\ & + k_{bb}^*(c_{r,s+1}^B - c_{r,s}^B) + k_{ba}^*c_{r+1,s}^A - k_{ab}^*c_{r,s}^B \end{aligned} \quad (32)$$

For $3 \leq n \leq N - 1$:

$$\begin{aligned} \frac{dc_{n,N-n}^A}{dt} = & k_{aa}c_{1,0}^Ac_{n-1,N-n}^A + k_{ba}c_{n-1,N-n}^Bc_{1,0}^A \\ & - k_{aa}^*c_{n,N-n}^A - k_{ba}^*c_{n,N-n}^A \end{aligned} \quad (33)$$

$$\begin{aligned} \frac{dc_{N-n,n}^B}{dt} = & k_{bb}c_{0,1}^Bc_{N-n,n-1}^B + k_{ab}c_{N-n,n-1}^Ac_{0,1}^B \\ & - k_{bb}^*c_{N-n,n}^B - k_{ab}^*c_{N-n,n}^B \end{aligned} \quad (34)$$

As remarked earlier, the complete reaction scheme must satisfy mass conservation in a closed system, implying that the mass variation rate must be strictly zero:

$$\begin{aligned} 0 = & 2\dot{c}_{1,1} + 3(\dot{c}_{2,1}^A + \dot{c}_{1,2}^B) + \sum_{n=1}^N n(\dot{c}_{n,0}^A + \dot{c}_{0,n}^B) \\ & + \sum_{n=2}^{N-1} (n+1)(\dot{c}_{1,n}^A + \dot{c}_{n,1}^B) + \sum_{n=2}^{N-2} (n+2)(\dot{c}_{2,n}^A + \dot{c}_{n,2}^B) \\ & + \sum_{r=3}^{N-1} \sum_{s=1}^{N-1} (r+s)(\dot{c}_{r,s}^A + \dot{c}_{s,r}^B), \end{aligned} \quad (35)$$

where the overdot stands for the time-derivative. The compliance with this constraint is an important and crucial check on the consistency of the numerical integration of the full set of differential equations eqn (14)–(34), which we monitor and confirm in all the simulations presented below. Analytically, this relation is satisfied by the rate equations.

As we see, there is one differential equation for each type of monomer and one for the heterodimer. The homopolymer set requires $2(N - 1)$ equations and the heteropolymer set a total of $N^2 - N - 1$ equations. The total number of kinetic differential

equations describing the whole system is $N(N + 1) - 1$, and is broken down into separate contributions as displayed in Table 1. Then, the total number of equations for describing the system as a function of maximum chain length N is:

$$\begin{aligned} \#eqs = & 5 + 2(N - 1) + 2(N - 2) + 2(N - 3) \\ & + (N^2 - 7N + 12) + 2(N - 3) = N(N + 1) - 1, \end{aligned} \quad (36)$$

as pointed out in ref. 28. From the computational point of view, the number of equations grows quadratically with the maximum chain length N .

3 Numerical results

We are interested in applying our copolymerization model to fit the experimental data measured by the Rehovot group, so our primary goal is to reproduce as closely as possible the details reported concerning the experiments on chiral amplification of oligopeptides. For this purpose, the first step is to determine the initial monomer concentrations to be employed in the simulations. The actual experiments were carried out for 0.5 mM solutions of monomers, thus we have employed for each case: (a) $R : S = 1 : 1$ which corresponds to an initial enantiomeric excess $ee_0 = 0\%$, so $c_{1,0}^A(0) = 0.25$ mM and $c_{0,1}^B(0) = 0.25$ mM; (b) $R : S = 4 : 6$ corresponding to $ee_0 = 20\%$, so $c_{1,0}^A(0) = 0.2$ mM and $c_{0,1}^B(0) = 0.3$ mM; (c) $R : S = 3 : 7$ which corresponds to $ee_0 = 40\%$, so $c_{1,0}^A(0) = 0.15$ mM and $c_{0,1}^B(0) = 0.35$ mM. The remainder of the initial concentrations (the dimers and on up) are taken to be zero. Next, we systematically search for the reaction rates leading to the best fit to the given data.

Different chemical model systems were used in the experiments: namely γ -stearyl-glutamic thioethyl ester (C_{18} -TE-Glu), N^ϵ -stearoyl-lysine thioethyl ester (C_{18} -TE-Lys), γ -stearyl-glutamic acid N -carboxyanhydride (C_{18} -Glu-NCA) and γ -stearyl-glutamic thioacid (C_{18} -thio-Glu), varying both their initial compositions and for various choices of catalyst. The composition of the oligopeptides formed was analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy (MALDI-TOF) with enantio-labeled samples. The experimental relative abundances of the oligopeptides were inferred from the ion intensity. These are the relative abundances that we aim to interpret *vis-a-vis* our copolymerization model.

Since only the experiments with racemic mixtures of the starting compounds required a catalyst, it is reasonable to expect that the racemic and the chiral enriched cases will follow different dynamics for a given model system. That is,

the presence or absence of a specific catalyst affects the rate constants, for a given chemical system. Firstly, we will find the reaction rates for the racemic case, and afterwards, those for the enriched chiral case, allowing us to compare both. The *a priori* nine free parameters we must set to run the numerical integrations are comprised of the four direct and the four inverse rate constants k_{aa} , k_{bb} , k_{ab} , k_{ba} , and k_{aa}^* , k_{bb}^* , k_{ab}^* , k_{ba}^* , plus the maximum polymer chain length, N . We set all the inverse reaction rates to a unique value, $k_{aa}^* = k_{bb}^* = k_{ab}^* = k_{ba}^* = 10^{-10}(\text{s}^{-1})$, implying an almost irreversible scheme, and we determine the remainder of the parameters from fitting the copolymerization model to the relative abundance data. This required numerical integration of the set of differential equations eqn (14)–(34) which we performed using the Mathematica program package. For each independent run we verified the compliance of the numerical results with the constraint in eqn (35), an imperative for any closed system.

Results from fitting the model to the data indicate that the maximum chain length N does not play a significant role, the Pearson product-moment correlation coefficient, r , remains the same for $N = 12, 14, 16, 18, 20$, so we will set $N = 12$ for all compounds and cases treated below. Since the number of

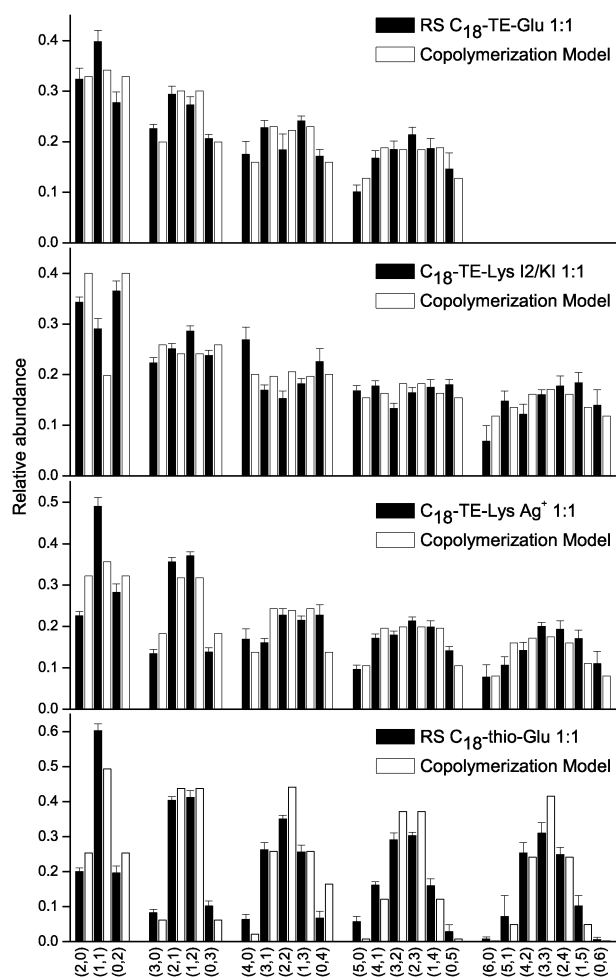


Fig. 2 Relative abundance *versus* number of repeat units (r,s) of the oligopeptides obtained from fitting the model (white) to the experimental data (black) from racemic mixtures $R : S = 1 : 1$ of monomers. The four chemical models are indicated by the insets.

independent equations scales as N^2 , this represents an important reduction in the computer time and the memory used. We note that one is free to scale out the dependence of one pair of reaction constants from the rate equations by a suitable redefinition of the time variable. Thus, without loss of generality, we set the cross inhibition rates equal to unity $k_{ab} = k_{ba} = 1 \text{ s}^{-1} \text{ mol}^{-1}$ and then search for the reaction rates $k_{aa} = k_{bb}$ leading to the best fits.

3.1 Racemic mixtures

In one set of experiments, the authors reported MALDI-TOF analysis of the oligopeptides formed at the air–water interface

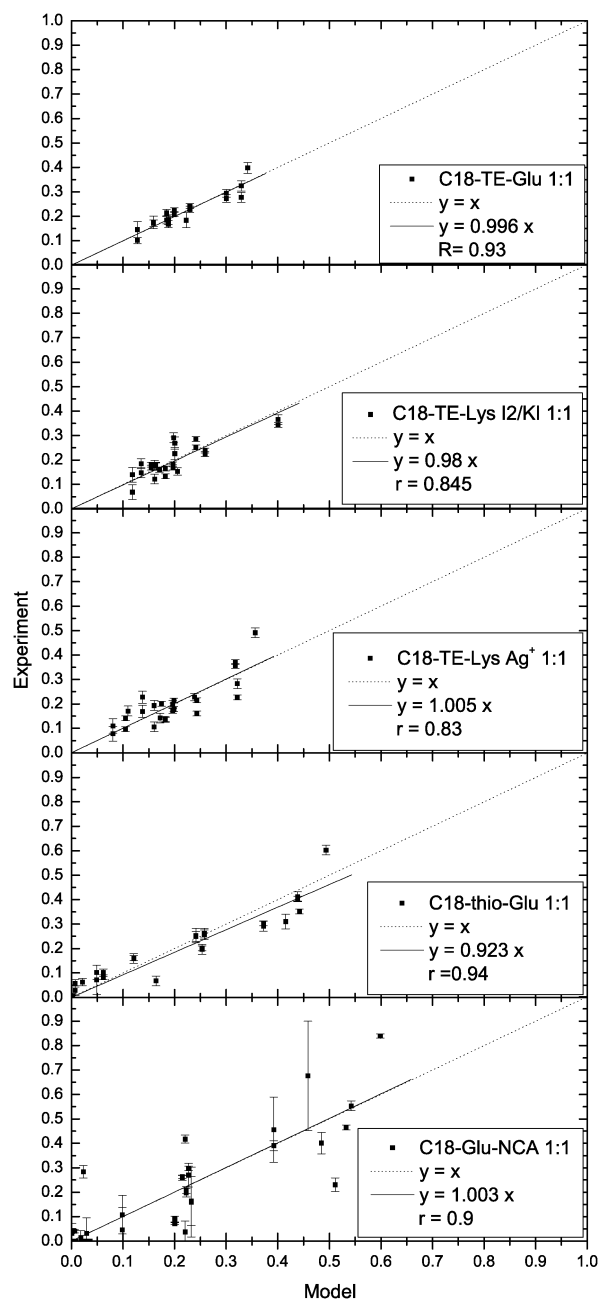


Fig. 3 Data correlations r from fitting the model to the data in the case of racemic mixtures of all the compounds employed. The chemical systems are indicated by the insets. The solid line represents the linear correlation between experimental data and numerical calculations.

from racemic mixtures $R : S = 1 : 1$ of the monomers for the various model systems and catalysts. We first fit the copolymerization model to this data.

The best correlation data for the racemic C_{18} -TE-Glu system with the I_2/KI catalyst are found for $k_{aa} = k_{bb} = 1.7 \text{ s}^{-1} \text{ mol}^{-1}$. In this case, the best fit is obtained for the time scale $t = 10^{11} \text{ s}$. Exactly by the same process, the best correlation data for the racemic C_{18} -TE-Lys system are found for $k_{aa} = k_{bb} = 2.3 \text{ s}^{-1} \text{ mol}^{-1}$ and for $k_{aa} = k_{bb} = 1.3 \text{ s}^{-1} \text{ mol}^{-1}$ when adding I_2/KI and $AgNO_3$ as catalysts, respectively. For the simulations here, we took the times $t = 10^{10} \text{ s}$ and $t = 10^{11} \text{ s}$ in the racemic cases with I_2/KI and $AgNO_3$ respectively. Finally, we fit our copolymerization model to the C_{18} -thio-Glu experimental relative abundances. The authors of the experiments affirmed that this compound undergoes a truly random polymerization, so fits from our model are expected to be slightly less satisfactory than those for the binomial distribution function. Setting the inverse reaction rates and the cross inhibition as indicated above, then the best correlation coefficients are found for $k_{aa} = k_{bb} = 0.4 \text{ s}^{-1} \text{ mol}^{-1}$. The instant or time-scale leading to these numerical values is $t = 10^{10} \text{ s}$.

The corresponding (experimental and numerical) relative abundances for the four compounds cited above corresponding to these values are shown in Fig. 2. The histograms show the relative abundance of each experimentally obtained oligopeptide compared to the best fit from our copolymerization model. We emphasize that we fit the model to the complete family of stereoisomer subgroups (global fit). The resulting data correlations are shown in Fig. 3 and Table 2, the latter gives a

Table 2 Comparative fits between the copolymerization model and the binomial distribution to the experimental relative abundances: racemic mixtures $R : S = 1 : 1$ of monomers of the four model systems as indicated in the leftmost column. Only in the case of C_{18} -thio-Glu does the binomial distribution give a better global fit than the copolymerization model: the latter system provides an experimental reference system for random polymerization²⁶

r	Copolymerization model					Bin.	
	Fits for each subgroup n					Global fit	Global fit
	Di	Tri	Tetra	Penta	Hexa		
C_{18} -TE-Glu	0.92	0.96	0.80	0.84	—	0.93	0.75
C_{18} -TE-Lys(I_2/KI)	0.9	-0.82	-0.11	-0.73	0.45	0.85	0.32
C_{18} -TE-Lys(Ag)	0.98	1	0.03	0.88	0.76	0.84	0.8
C_{18} -thio-Glu	1	1	1	0.98	0.97	0.95	0.98

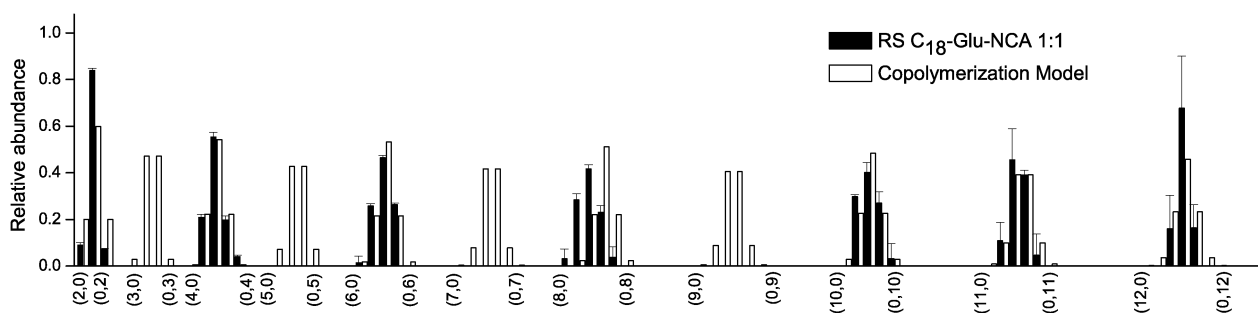


Fig. 4 The C_{18} -Glu-NCA system with catalyst $Ni(CH_3CO_2)_2$: relative abundance of the oligopeptides obtained from fitting the model (white) to the experimental data (black) from racemic mixtures of monomers. Compare Fig. 4A of ref. 26.

detailed comparison of the best fits between individual subfamilies and the overall global fit.

In the case of the C_{18} -Glu-NCA with catalyst $Ni(CH_3CO_2)_2$, the best fit is obtained for $k_{aa} = k_{bb} = 0.2 \text{ s}^{-1} \text{ mol}^{-1}$. Results for the corresponding relative abundances are shown in Fig. 4 and the correlation from fitting is displayed in the bottom frame of Fig. 3 and Table 3. Not all subfamily data sets are reported in the experimental paper;²⁵ here we use the fitted model to the partial data set to predict or fill in these missing subfamily data. Numerical results for the racemic case have been found for $t = 10^{10} \text{ s}$.

3.2 Chirally enriched mixtures

In a second set of experiments, the authors reported MALDI-TOF analysis of the oligopeptides formed at the air-water interface from non-racemic mixtures of the monomers for the same model systems. No catalysts were employed there. We next consider fits of our model to these data sets.

The best correlation factors for both chirally enriched mixture cases (20% and 40% excesses) in the case of the C_{18} -TE-Glu system are found for the same rates, that is for $k_{aa} = k_{bb} = 2 \text{ s}^{-1} \text{ mol}^{-1}$. The results for these values are shown in Table 4. In Fig. 5 we display the relative abundances of the homochiral oligopeptides and in Table 5 both the calculated and experimental enantiomeric excesses for the 4 : 6 and 3 : 7 ($R : S$) mixtures. In Fig. 6 we show the data correlation. Numerical results for the non-racemic case have been found for the time scale $t = 10^{11} \text{ s}$.

For the chiral mixtures of C_{18} -TE-Lys we found the best fits for the dynamics corresponding to $k_{aa} = k_{bb} = 2.5 \text{ s}^{-1} \text{ mol}^{-1}$. The results for these values are shown in Table 6. The relative abundances results for these values are shown in Fig. 7 and the enantiomeric excesses obtained for 4 : 6 and 3 : 7 ($R : S$) mixtures are presented in Table 7. In Fig. 8 the data correlation is shown. For the simulations here, we took the instants $t = 10^{10} \text{ s}$ and $t = 10^{11} \text{ s}$ in the racemic cases with I_2/KI and $AgNO_3$ respectively, and $t = 10^{10} \text{ s}$ for the chirally enriched mixtures.

In the case of nonracemic C_{18} -thio-Glu, the best correlation coefficients are found for the same values of the reaction rates that we found in the racemic case, namely for $k_{aa} = k_{bb} = 0.4 \text{ s}^{-1} \text{ mol}^{-1}$. Results for the chiral cases are shown in Table 8. As to be expected and as shown in the table, the correlation factors for the global fit to the binomial distribution function are

Table 3 Comparative fits between the copolymerization model and binomial distribution to the experimental relative abundances for the racemic compositions ($R : S = 1 : 1$) of the C_{18} -Glu-NCA system

r	Copolymerization model												Bin.	
	Fits for each subgroup n												Global fit	Global fit
	Di	Tri	Tetra	Penta	Hexa	Hepta	Octa	Nona	Deca	Endeca	Dodeca	Global fit		
C_{18} -Glu-NCA	1	—	1	—	0.98	—	0.98	—	0.97	1	0.95	0.96	0.75	

Table 4 Comparative fits between the copolymerization model and the binomial distribution to the experimental relative abundances measured for non-racemic mixtures of C_{18} -TE-Glu

r	Copolymerization model						Binomial
	Fits for each subgroup n						Global fit
	Di	Tri	Tetra	Penta	Hexa	Global fit	
($R : S$) 4 : 6	0.86	0.89	0.93	0.99	—	0.94	0.75
($R : S$) 3 : 7	0.95	0.94	0.96	0.99	0.99	0.95	0.85

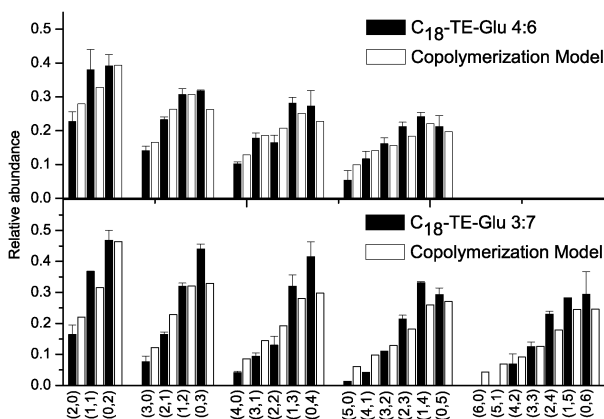


Fig. 5 Relative abundance versus number of repeat units (r,s) of the oligopeptides obtained from fitting the model (white) to the experimental data (black) from non-racemic mixtures of monomers for the C_{18} -TE-Glu system.

Table 5 Enantiomeric excesses ee : numerical results from the copolymerization model (experimental data) for the relative abundances of the homochiral oligopeptides for the C_{18} -TE-Glu system

$ee(\%)$	Di	Tri	Tetra	Penta	Hexa
($R : S$) 4 : 6	18 (26)	24 (39)	30 (46)	35 (59)	—
($R : S$) 3 : 7	37 (48)	48 (71)	57 (82)	66 (92)	73 (>99.8)

slightly better than those for any simulation we could perform with the copolymerization model, so we reconfirm what was claimed by the authors of the experimental work: namely that the C_{18} -thio-Glu system polymerizes randomly. In Fig. 9 the relative abundances of the oligopeptides are shown. The data correlation is shown in Fig. 10.

The best fits for both chirally enriched mixture cases (20% and 40% excesses) in the case of C_{18} -Glu-NCA are found for the same dynamics, that is $k_{aa} = k_{bb} = 3.8 \text{ s}^{-1} \text{ mol}^{-1}$. The results for these values are shown in Table 9. In Fig. 11 we compare the best fit against the experimentally obtained

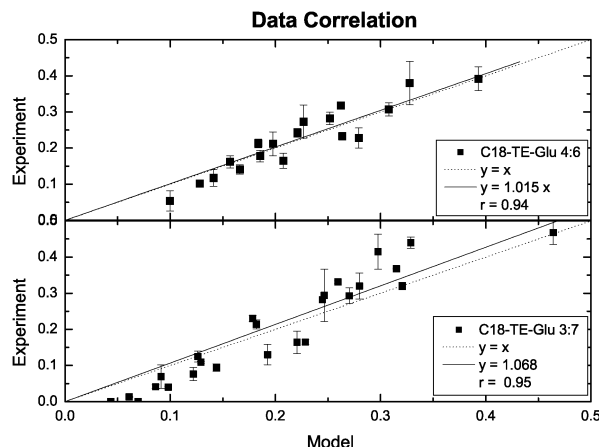


Fig. 6 Results from fitting the model to the experimental data: non-racemic mixtures of C_{18} -TE-Glu. The solid line represents the linear correlation between experimental and numerical data obtained from fitting. The dotted line has a slope equal to unity.

Table 6 Results for the copolymerization model and experimental data correlations for non-racemic mixtures of C_{18} -TE-Lys

r	Copolymerization model							Binomial
	Fits for each subgroup n							Global fit
	Di	Tri	Tetra	Penta	Hexa	Hepta	Global fit	
($R : S$) 4 : 6	0.78	1	0.87	0.90	0.84	0.97	0.89	0.46
($R : S$) 3 : 7	0.93	1	0.95	0.97	0.99	—	0.94	0.65

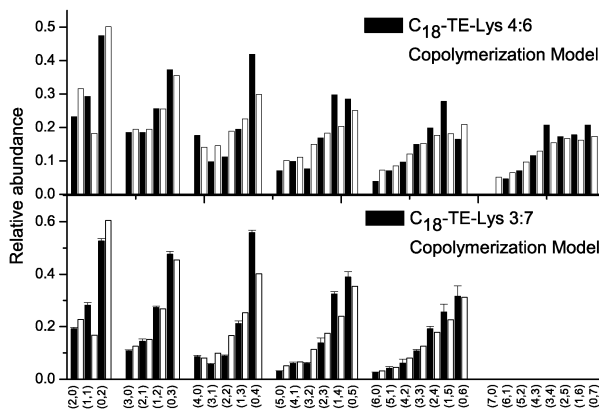


Fig. 7 Relative abundance versus number of repeat units (r,s) of the oligopeptides obtained from fitting the model (white) to the experimental data (black) from non-racemic mixtures of monomers of C_{18} -TE-Lys.

Table 7 Enantiomeric excesses: numerical results from the copolymerization model (experimental data) for the relative abundances of the homochiral oligopeptides for C₁₈–TE–Lys

ee(%)	Di	Tri	Tetra	Penta	Hexa	Hepta
(R : S) 4 : 6	23 (34)	30 (34)	36 (41)	42 (60)	49 (62)	54 (>99.8)
(R : S) 3 : 7	45 (46)	57 (63)	66 (73)	75 (85)	81 (86)	—

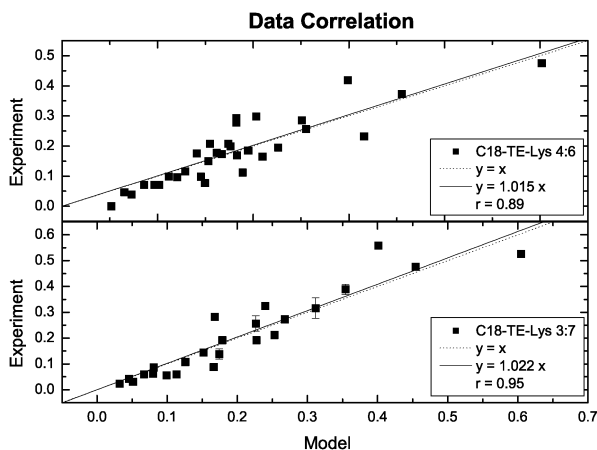


Fig. 8 Results from fitting the model to the experimental data. Chiral mixtures of C₁₈–TE–Lys. The solid line represents the linear correlation between the experimental and numerical data obtained from fitting. The dotted line has a slope equal to unity.

Table 8 Results for the copolymerization model and experimental data correlations for non-racemic mixtures of C₁₈–thio–Glu

<i>r</i>	Copolymerization model							Binomial
	Di	Tri	Tetra	Penta	Hexa	Hepta	Global fit	Global fit
(R : S) 4 : 6	0.93	0.98	0.93	0.92	0.92	0.91	0.91	0.93
(R : S) 3 : 7	0.89	1	0.99	0.99	0.98	—	0.96	0.97

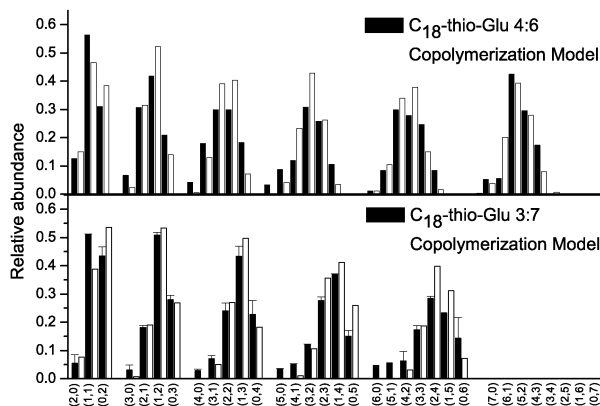


Fig. 9 Relative abundances *versus* number of repeat units (*r,s*) of the oligopeptides obtained from fitting the model (white) to the experimental data (black) for the non-racemic mixtures of C₁₈–thio–Glu.

relative abundances of the oligopeptides. The corresponding data correlation is shown in Fig. 12. Numerical results for the racemic case have been found for $t = 10^{11}$ s.

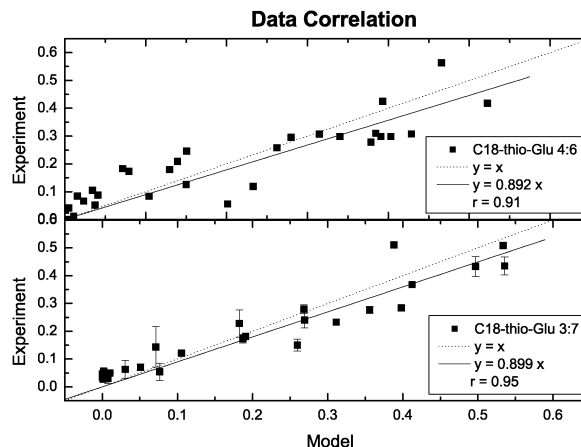


Fig. 10 Results from fitting the model to the experimental data for non-racemic mixtures of the C₁₈–thio–Glu system. The solid line represents the linear correlation between experimental and numerical data obtained from fitting. The dotted line has a slope equal to unity.

4 Conclusions

The overall scheme for the chiral amplification process leading to the experimental data investigated here involves a self-assembly step followed by a lattice-controlled polymerization.^{25,26} It is this subsequent polymerization which is the prime focus of this paper. The authors of the experimental work stress that it is not at all straightforward to actually establish the *correlation* between the packing arrangements of the crystallites and the composition of the diastereoisomeric products that result therefrom. Therefore, our task here was to fit the outcome of these latter steps assuming an *effective* copolymerization scheme. The term “effective” simply means that the putative complicated correlations and interplay between the 2D crystallite phases at the air–water interface and the polymerization reaction pathways that depend on the microscopic packing arrangements within the crystals are treated here with a simple model. In this regard, our model can be regarded as a “course-grained” description of the overall process in that the microscopic details (the structures of the crystalline phases) are not resolved, but that the end-result or net effect of the pathways afforded by the crystallites can be summarized by the polymerization scheme as depicted graphically in Fig. 1.

The model as introduced is defined for fully reversible reactions and this implies that some of the reaction rates must obey a corresponding constraint as dictated by microreversibility. Thus the model is appropriate for closed systems under thermodynamic control. For the numerical fits themselves, we found that all the reverse reaction rates could be set to rather small values, and this is in consonance with experimentally observed irreversible condensation. Thus for the present purposes, the copolymerization model is practically irreversible. The values for the forward rates of adding the same chirality monomer to the end of the growing chain are found to be greater than those for addition of a wrong chirality monomer: that is, $k_{aa} = k_{bb} > k_{ab} = k_{ba} = 1$ (except of course for the model system C₁₈–thio–Glu serving as a reference for random polymerization).

Other closed systems that lead to copolymers could, in principle, be fit with our model. If for example k_{aa} and k_{bb} had different magnitudes, this would imply that an underlying chiral bias is operative either in the polymerization or in the

Table 9 Results for the copolymerization model and experimental data correlations for C₁₈–Glu–NCA. The global fit from the binomial distribution is shown for comparison

<i>r</i>	Copolymerization model									Binomial	
	Fits for each subgroup <i>n</i>									Global fit	Global fit
	Di	Tri	Tetra	Penta	Hexa	Hepta	Octa	Nona	Deca		
(<i>R</i> : <i>S</i>) 4 : 6	−0.79	0.9	0.63	0.74	0.95	0.89	0.77	0.86	—	0.68	0.11
(<i>R</i> : <i>S</i>) 3 : 7	−0.33	0.79	0.81	0.76	0.86	0.96	0.75	0.83	0.89	0.75	0.38

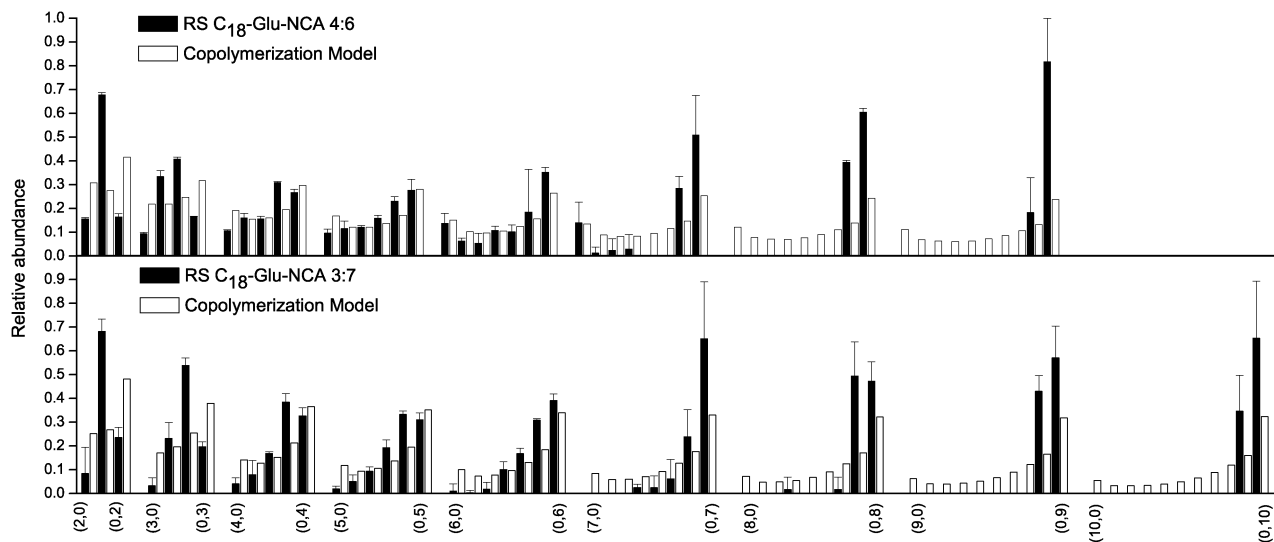


Fig. 11 Relative abundances for the non-racemic mixtures of C₁₈–thio–Glu. The experimental data set (black) is incomplete, we have used our model to fill in the missing portions of the histogram (white).

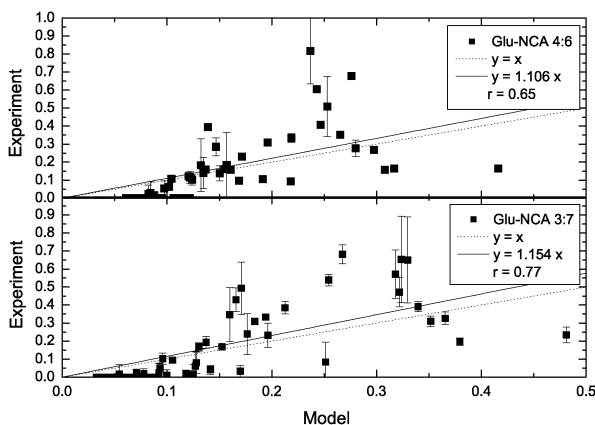


Fig. 12 Results from fitting the model to the experimental data for the C₁₈–Glu–NCA system. The solid line represents the linear correlation between experimental and numerical data obtained from fitting. The dotted line has a slope equal to unity.

prior formation of the two crystallites that control the polymerization. This bias could affect the packing arrangements of the crystal monomers and the reaction pathways taken within each crystallite phase. Since however our model is effective, as explained earlier, we would not be able to say whether the chiral bias is in the polymerization or in the structure of the crystallites that control the polymerization. Nevertheless, this bias in k_{aa} being different from k_{bb} would result in favoring the attachment of say, an *S* to an *S* over the attachment of an *R*

to an *R*, and this feature would show clearly up in the relative abundances.

Another positive feature of the model is the robustness of the fits with respect to differing initial imbalances of the enantiomers. That is, for a given chemical model (including catalyst, if any) the values of the fitted rates do not depend on the initial enantiomeric excesses of the monomers. If our rate constants are viewed as effective, that is, implicitly involving the different chemical properties of the racemic and enantiomorphous crystallite phases, then this feature suggests that the packing arrangements and reaction pathways in the solid-state do not depend (or only weakly) on the magnitude of these imbalances.

The Pearson product-moment correlation coefficient r between experimental and numerical data is greater for the copolymerization model than for the binomial distribution, except for the C₁₈–thio–Glu, which truly polymerizes randomly. The correlation between calculated and experimental relative abundances is also greater for the initially non-racemic situations, and the higher the initially chiral enrichment of the mixture is, the better the copolymerization model reproduces the chemical data. The results obtained here lead us to affirm that the model systems considered all undergo a non-random polymerization, as was asserted by the authors of the experiments.^{25,26}

The model also qualitatively reproduces the behavior of the enantiomeric excess *ee*, its increase with the length of the chains and the enhancement of the *ee* of the corresponding initial mixture of monomers. All this, in spite of the complexity

of the factors that affect the reactivity within the experimental two-phase system, *i.e.*, the microscopic crystallite packing arrangements and the possible reaction pathways within these 2D crystallites. In conclusion, we may therefore assert that our simple scheme does provide an accurate course-grained description of the lattice-controlled polymerization reported in ref. 25 and 26.

Acknowledgements

We are grateful to Meir Lahav for providing us with the experimental data and for many helpful discussions and correspondence. CB has a Calvo-Rodés predoctoral scholarship from the Instituto Nacional de Técnica Aeroespacial (INTA) and the research of DH is supported in part by the grant AYA2009-13920-C02-01 from the Ministerio de Ciencia e Innovación (Spain) and forms part of the COST Action CM0703 “Systems Chemistry”.

References

- 1 G. Joyce, G. Visser, C. van Boeckel, J. van Boom, L. Orgel and J. van Westrenen, *Nature*, 1984, **310**, 602–604.
- 2 V. Avetisov and V. Goldanskii, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 11435–11442.
- 3 D. Kondepudi and K. Asakura, *Acc. Chem. Res.*, 2001, **34**, 946–954.
- 4 P. Cintas, *Angew. Chem., Int. Ed.*, 2002, **41**, 1139–1145.
- 5 C. Huber, W. Eisenreich, S. Hecht and G. Wächtershäuser, *Science*, 2003, **301**, 938–940.
- 6 L. Leman, L. Orgel and M. Ghadiri, *Science*, 2004, **306**, 283–286.
- 7 C. Huber and G. Wächtershäuser, *Science*, 1997, **276**, 245–247.
- 8 C. deDuve, *Blueprint for a cell: The nature and origin of life*, Neil Patterson Publisher, Burlington, North Carolina, 1st edn, 1991.
- 9 R. Pascal, L. Boiteau and A. Commeyras, *Top. Curr. Chem.*, 2005, **259**, 69–122.
- 10 R. Lundberg and P. Doty, *J. Am. Chem. Soc.*, 1957, **79**, 3961–3972.
- 11 M. Idelson and E. Blout, *J. Am. Chem. Soc.*, 1957, **79**, 3948–3955.
- 12 T. Akaike and S. Inoue, *Biopolymers*, 1976, **15**, 1863–1868.
- 13 N. Blair and W. Bonner, *Origins Life Evol. Biosphere*, 1980, **10**, 255–263.
- 14 N. Blair and W. Bonner, *Origins Life Evol. Biosphere*, 1981, **11**, 331–335.
- 15 H. Kricheldorf, *Angew. Chem., Int. Ed.*, 2006, **45**, 5752–5784.
- 16 K. Ehler and L. Orgel, *Biochem. Biophys. Acta*, 1976, **434**, 233–243.
- 17 A. Hill, C. Bohler and L. Orgel, *Origins Life Evol. Biosphere*, 1998, **28**, 235–243.
- 18 A. Brack, *Origins Life Evol. Biosphere*, 1987, **17**, 367–379.
- 19 H. Kanazawa, *Polymer*, 1992, **33**, 2557–2566.
- 20 H. Kanazawa and Y. Ohashi, *Mol. Cryst. Liq. Cryst.*, 1996, **277**, 45–54.
- 21 T. Hitz, M. Blocher, P. Walde and P. Luisi, *Macromolecules*, 2001, **34**, 2443–2449.
- 22 T. Hitz and P. Luisi, *Helv. Chim. Acta*, 2002, **85**, 3975–3983.
- 23 T. Hitz and P. Luisi, *Helv. Chim. Acta*, 2003, **86**, 1423–1434.
- 24 M. Blocher, T. Hitz and P. Luisi, *Helv. Chim. Acta*, 2001, **84**, 842–847.
- 25 H. Zepik, E. Shavit, M. Tang, T. Jensen, K. Kjaer, G. Bolbach, L. Leiserowitz, I. Weissbuch and M. Lahav, *Science*, 2002, **295**, 1266–1269.
- 26 I. Weissbuch, H. Zepik, G. Bolbach, E. Shavit, M. Tang, T. Jensen, K. Kjaer, L. Leiserowitz and M. Lahav, *Chem.–Eur. J.*, 2003, **9**, 1782–1794.
- 27 J. A. D. Wattis and P. V. Coveney, *J. Phys. Chem. B*, 2007, **111**, 9546–9562.
- 28 C. Blanco and D. Hochberg, *Phys. Chem. Chem. Phys.*, 2011, **13**, 839–849.

Cite this: *Chem. Commun.*, 2012, **48**, 3659–3661

www.rsc.org/chemcomm

Induced mirror symmetry breaking *via* template-controlled copolymerization: theoretical insights†‡

Celia Blanco and David Hochberg*

Received 23rd December 2011, Accepted 20th February 2012

DOI: 10.1039/c2cc18045f

A chemical equilibrium model of template-controlled copolymerization is presented for describing the outcome of the experimental induced desymmetrization scenarios recently proposed by Lahav and co-workers.

It is an empirical fact that mirror symmetry is broken in all known biological systems, where processes crucial for life such as replication, imply chiral supramolecular structures, sharing the same chiral sign (homochirality). These chiral structures are proteins, composed of amino acids almost exclusively found as the left-handed enantiomers (*S*), also DNA, and RNA polymers and sugars with chiral building blocks composed by right-handed (*R*) monocarbohydrates.

One scenario for the transition from prebiotic racemic chemistry to chiral biology suggests that homochiral peptides must have appeared before the onset of the primeval enzymes.^{1–5} However, the polymerization of racemic mixtures (1 : 1 proportions) of monomers in ideal solutions typically yields chains composed of random sequences of both the left- and right-handed repeat units following a binomial distribution.⁶ This statistical problem has been overcome recently by the experimental demonstration of the generation of amphiphilic peptides of homochiral sequence, that is, of a single chirality, from racemic compositions. This route consists of two steps: (1) the formation of racemic parallel or anti-parallel β -sheets either in aqueous solution or in 3-D crystals⁷ during the polymerization of racemic hydrophobic α -amino acids followed by (2) an enantioselective controlled polymerization reaction^{8–14} (Fig. 1). This process leads to racemic or mirror-symmetric mixtures of isotactic oligopeptides where the chains are composed from amino acid residues of a single handedness. Furthermore,



Fig. 1 The scheme proposed in ref. 7 leading to regio-enantioselection within racemic β -sheet templates.

Department of Molecular Evolution, Centro de Astrobiología (CSIC-INTA), Ctra. Ajalvir Km. 4, 28850 Torrejón de Ardoz, Madrid, Spain. E-mail: blancodtc@cab.inta-csic.es, hochbergd@cab.inta-csic.es

† This article is part of the ChemComm 'Chirality' web themed issue.

‡ Electronic supplementary information (ESI) available. See DOI: 10.1039/c2cc18045f

when racemic mixtures of different amino acid species were polymerized, isotactic co-peptides of homochiral sequence were generated. Here a host or majority species (R_0, S_0), together with a given number m of minority amino acid species ($R_1, S_1, R_2, S_2, \dots, R_m, S_m$) (supplied with lesser abundance) were employed. The guest (*S*) and (*R*) molecules are enantioselectively incorporated into the chains of the (*S*) and (*R*) peptides, respectively, however the former are *stochastically* distributed within the homochiral chains. As a combined result of these two effects, the sequence of the co-peptide *S* and *R* chains will differ from each other, resulting in non-racemic mixtures of co-peptide polymer chains: *non-enantiomeric pairs* of chains are thus formed. By considering the sequences of these peptide chains, a statistical departure from the racemic composition of the library of the peptide chains is created which varies with chain length N and with the relative concentrations of the host/guest monomers used in the polymerization.^{9,10} The mechanism has some features in common with the scenarios proposed by Green and Garetz,¹⁵ Eschenmoser *et al.*¹⁶ and Siegel¹⁷ in which a limited supply of material results in a stochastic mirror symmetry breaking process.

To address the general scenario for the generation of libraries of diastereoisomeric mixtures of peptides in accord with that proposed in ref. 9, consider a model with a host amino acid species and m guest amino acids. We assume as given the prior formation of the initial templates or β -sheets, and are concerned exclusively with the subsequent random polymerization reactions (step (2)). The underlying nonlinear template control is implicit throughout the discussion.

We consider stepwise additions and dissociations of single monomers from one end of the (co)polymer chain, considered as a strand within the β -sheet. It is reasonable to regard the β -sheet in equilibrium with the free monomer pool.¹⁸ §

From detailed balance, each individual monomer attachment or dissociation reaction is in equilibrium. This holds for closed equilibrium systems in which the free monomers are depleted/replenished by the templated polymerization. Then we can compute the equilibrium concentrations of all the (co)-polymers in terms of equilibrium constants K_i and the free monomer concentrations. The equilibrium concentration of an *S*-type copolymer chain of length $n_0 + n_1 + n_2 + \dots + n_m = N$ made up of n_j molecules S_j is given by $p_{n_0, n_1, \dots, n_m}^S = (K_0 s_0)^{n_0} (K_1 s_1)^{n_1} \dots (K_m s_m)^{n_m} / K_0$, where $s_j = [S_j]$.¹⁹ Similarly for the concentration of an *R*-type copolymer

chain of length $n'_0 + n'_1 + n'_2 + \dots + n'_m = N$ made up of n'_j molecules R_j : $p_{n'_0, n'_1, \dots, n'_m}^R = (K_0 r_0)^{n'_0} (K_1 r_1)^{n'_1} \dots (K_m r_m)^{n'_m} / K_0$, where $r_j = [R_j]$.

The number of different S -type copolymers of length l with n_j molecules of type S_j is given by the multinomial coefficient. Hence the total concentration of the S -type copolymers of length l is given by

$$p_l^S = \sum_{n_0+n_1+\dots+n_m=l} \binom{l}{n_0, n_1, \dots, n_m} p_{n_0, n_1, \dots, n_m}^S \quad (1)$$

$$= \frac{1}{K_0} (K_0 s_0 + K_1 s_1 + \dots + K_m s_m)^l,$$

which follows from the multinomial theorem.²⁰ We calculate the number of each type S_j of S -monomer present in the S -copolymer of length equal to l , for any $0 \leq j \leq m$:

$$s_j(p_l^S) = \sum_{n_0+n_1+\dots+n_m=l} \binom{l}{n_0, n_1, \dots, n_m} n_j p_{n_0, n_1, \dots, n_m}^S$$

$$= s_j \frac{\partial}{\partial s_j} p_l^S = \frac{K_j}{K_0} s_j l (K_0 s_0 + K_1 s_1 + \dots + K_m s_m)^{l-1}. \quad (2)$$

Then we need to know the total amount of the S -type monomers bound within the S -type copolymers, from the dimer on up to a maximum chain length N . Using eqn (2) for the j th type of amino acid, this is given by

$$s_j(p_{\text{Tot}}^S) = \sum_{l=2}^N s_j(p_l^S) \rightarrow \frac{K_j}{K_0} s_j \frac{a(2-a)}{(1-a)^2}, \quad (3)$$

the final expression holds in the limit $N \rightarrow \infty$ provided that $a = (K_0 s_0 + K_1 s_1 + \dots + K_m s_m) < 1$. This must be the case, otherwise the system would contain an infinite number of molecules.¹⁹ Similar considerations hold for the R -sector, and the total amount of R monomers inside R type copolymers for the j th amino acid is given by $r_j(p_{\text{Tot}}^R) = \frac{K_j}{K_0} r_j \frac{b(2-b)}{(1-b)^2}$ where $b = (K_0 r_0 + K_1 r_1 + \dots + K_m r_m) < 1$. From this we obtain the mass balance equations which hold for both enantiomers of the host and guest amino acids, and is our key result:

$$s_j + \frac{K_j}{K_0} s_j \frac{a(2-a)}{(1-a)^2} = s_{j\text{tot}}, \quad r_j + \frac{K_j}{K_0} r_j \frac{b(2-b)}{(1-b)^2} = r_{j\text{tot}}. \quad (4)$$

These equations express the fact that each type of enantiomer is either free, or else is bound inside a (co)polymer strand within the template.

The problem then consists in the following: given the total concentrations of all the $m + 1$ enantiomers $\{s_{j\text{tot}}, r_{j\text{tot}}\}_{j=0}^m$, and the K_i we calculate the free monomer concentrations $\{s_j, r_j\}_{j=0}^m$ from solving eqn. (4). Denote by $s_{0\text{tot}} + \dots + s_{m\text{tot}} + r_{0\text{tot}} + \dots + r_{m\text{tot}} = c_{\text{tot}}$ the total system concentration. From the solutions we can calculate *e.g.*, the equilibrium concentrations of homochiral copolymers of any specific sequence or composition as well as the resultant enantiomeric excess for homochiral chains of length l composed of the host (majority) amino acid: $ee_l = \frac{(r_0)^l - (s_0)^l}{(r_0)^l + (s_0)^l}$. When there are no guest amino acids, *i.e.*, for $m = 0$, and when the majority species is supplied in racemic

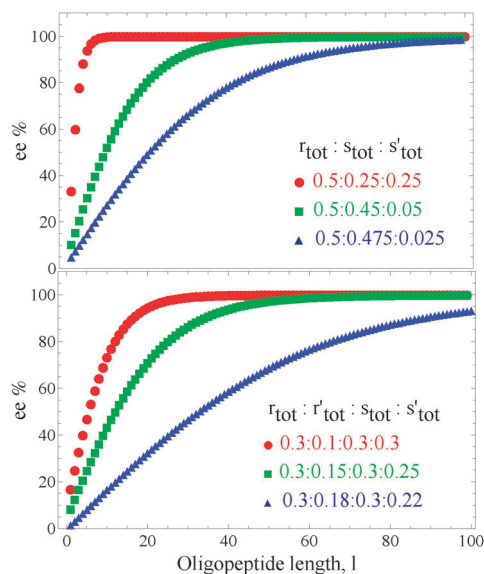


Fig. 2 Calculated ee values versus chain length l from solving eqn (4). (Top) Non-racemic host $r_{\text{tot}} > s_{\text{tot}}$ and one guest amino acid s'_{tot} ($m = 1$) and three monomer starting compositions (in moles) $r_{\text{tot}} : s_{\text{tot}} : s'_{\text{tot}} = 0.5 : 0.25 : 0.25$ (filled circles), $0.5 : 0.45 : 0.05$ (squares) and $0.5 : 0.475 : 0.025$ (triangles) for the equilibrium constant $K = 1 \text{ M}^{-1}$ and the total monomer concentration $c_{\text{tot}} = 1 \text{ M}$. Compare to Fig. 13 of ref. 9. (Bottom) Racemic host $r_{\text{tot}} = s_{\text{tot}}$ and $m = 1$ guest $r'_{\text{tot}}, s'_{\text{tot}}$. Starting compositions $r_{\text{tot}} : r'_{\text{tot}} : s_{\text{tot}} : s'_{\text{tot}} = 0.3 : 0.1 : 0.3 : 0.3$ (filled circles), $0.3 : 0.15 : 0.3 : 0.25$ (squares) and $0.3 : 0.18 : 0.3 : 0.22$ (triangles) for four monomers.

proportions $s_{0\text{tot}} : r_{0\text{tot}} = 1 : 1$, then ee_l must be zero: there will be no mirror symmetry breaking. So we turn to the scenario of ref. 9 and consider the influence of a single guest species, $m = 1$ being sufficient for our purposes.

We first use our mass balance equations to calculate ee_l for the same initial compositions of the monomers as reported in ref. 9. This is shown in top of Fig. 2. We consider a single equilibrium constant $K_0 = K_1 = K = 1 \text{ M}^{-1}$ for sake of simplicity, and the total system concentration, $c_{\text{tot}} = 1 \text{ M}$. The enantiomeric excess increases when increasing the amount of guest species s'_{tot} , obtaining a maximal symmetry breaking for the case shown with equal amounts of majority and minority S -molecules: $s_{\text{tot}} = s'_{\text{tot}}$. In the limit as $s'_{\text{tot}} \rightarrow 0$ we tend towards a racemic situation, so decreasing the amount of the minority or guest species is equivalent to approaching the racemic state, manifested through ever smaller values of ee_l for fixed l (top to bottom sequence of curves). The ee_l increases monotonically with the chain length l in all cases. The behavior of the ee_l demonstrates quite well the induced symmetry breaking mechanism proposed in ref. 9.

The solutions of the mass balance equations (4) can be used to evaluate the average chain lengths as functions of initial monomer compositions and the equilibrium constants. The average chain lengths of the S -type copolymers $\langle l_S \rangle$, composed of random sequences of the S_j type monomers, and that of the R -type copolymers $\langle l_R \rangle$ composed of random sequences of the R_j type monomers, are derived in the ESI.† Results for the $m = 1$ three monomer cases are shown in Table SI, ESI.† There is a marked increase in the average chain length when increasing K , we moreover observe how the average

chain length corresponding to each monomer species increases when increasing its own starting proportion. In the case of additives of only one handedness (three monomer case) and for the different compositions considered ($r_{\text{tot}} : s_{\text{tot}} : s'_{\text{tot}} = 0.5 : 0.25 : 0.25$, $0.5 : 0.45 : 0.05$, $0.5 : 0.475 : 0.025$) the average chain length for the *S*-type copolymers and the *R*-type polymers will be the same. This follows since K is the same for both monomer types and the amount of *S*-type and *R*-type molecules in the starting compositions is the same, $r_{\text{tot}} = s_{\text{tot}} + s'_{\text{tot}}$, so the average chain length must be the same: $\langle l_S \rangle = \langle l_R \rangle$.

By a further example, we carry out an analysis for the case of one guest $m = 1$ and all four enantiomers, treating a majority species *R*, *S* in strictly racemic proportions and a single guest amino acid R' , S' in various relative proportions. We solve eqn (4) and then calculate ee_l for the different chain lengths l for three different starting monomer compositions. In Fig. 2 (bottom) we show the results obtained from calculating ee_l for $K = 1 \text{ M}^{-1}$ and $c_{\text{tot}} = 1 \text{ M}$. The behavior is qualitatively similar to that previously commented, the greater the relative disproportionation of the minority species r'_{tot} , s'_{tot} , the greater is the enantiomeric excess. Values for the average chain lengths are calculated for four molecules, with the abundances $r_{\text{tot}} : r'_{\text{tot}} : s_{\text{tot}} : s'_{\text{tot}} = 0.3 : 0.1 : 0.3 : 0.3$ and $r_{\text{tot}} : r'_{\text{tot}} : s_{\text{tot}} : s'_{\text{tot}} = 0.3 : 0.14 : 0.3 : 0.26$, and are displayed in Table SII in ESI†, where other choices for the K_i and c_{tot} are employed (see Tables SIII to SVI, ESI†).

In summary, we consider a multinomial sample space for the distribution of equilibrium concentrations of homochiral copolymers formed *via* template control. We deduce mass balance equations for the enantiomers of the individual amino acid species, and their solutions are used to evaluate the sequence-dependent copolymer concentrations, in terms of the total species concentrations. Measurable quantities signalling the degree of mirror symmetry breaking such as the ee and average chain lengths are evaluated. This approach provides a quantitative basis for the template-controlled induced desymmetrization mechanisms advocated by Lahav and co-workers.^{8–14}

We are indebted to Meir Lahav for suggesting a mathematical approach to this problem. CB has a Calvo Rodés scholarship

from INTA. DH acknowledges a grant AYA2009-13920-C02-01 from the MICINN and forms part of the COST Action CM0703 “Systems Chemistry”.

Notes and references

§ Ref. 18 reports a stochastic simulation of two concurrent processes: (1) an irreversible condensation of activated amino acids and (2) reversible formation of racemic β -sheets of alternating homochiral strands, treated as a one-dimensional problem. These architectures lead to the formation of chiral peptides whose isotacticity increases with length.

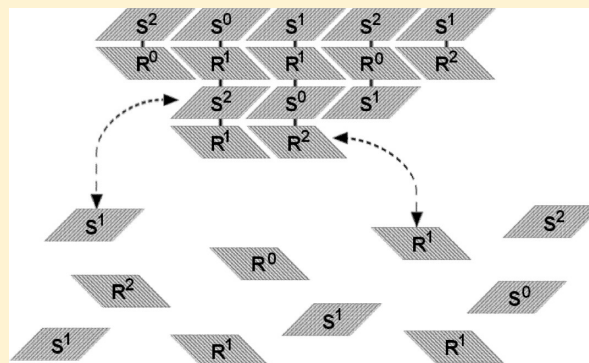
- 1 J. Bada and S. Miller, *Biosystems*, 1987, **20**, 21.
- 2 V. Avetisov, V. Goldanskii and V. Kuzmin, *Dokl. Akad. Nauk SSSR*, 1985, **115**, 282.
- 3 V. Goldanskii, V. Avetisov and V. Kuzmin, *FEBS Lett.*, 1986, **207**, 181.
- 4 V. Avetisov and V. Goldanskii, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 11435.
- 5 L. Orgel, *Nature*, 1992, **358**, 203.
- 6 A. Guijarro and M. Yus, *The Origin of Chirality in the Molecules of Life*, RSC Publishing, Cambridge, 1st edn, 2009.
- 7 I. Weissbuch, R. Illos, G. Bolbach and M. Lahav, *Acc. Chem. Res.*, 2009, **42**, 1128.
- 8 H. Zepik, E. Shavit, M. Tang, T. Jensen, K. Kjaer, G. Bolbach, L. Leiserowitz, I. Weissbuch and M. Lahav, *Science*, 2002, **295**, 1266.
- 9 J. Nery, G. Bolbach, I. Weissbuch and M. Lahav, *Chem.–Eur. J.*, 2005, **11**, 3039.
- 10 J. Nery, R. Eliash, G. Bolbach, I. Weissbuch and M. Lahav, *Chirality*, 2007, **19**, 612.
- 11 I. Rubinstein, R. Eliash, G. Bolbach, I. Weissbuch and M. Lahav, *Angew. Chem., Int. Ed.*, 2007, **46**, 3710.
- 12 I. Rubinstein, G. Clodic, G. Bolbach, I. Weissbuch and M. Lahav, *Chem.–Eur. J.*, 2008, **14**, 10999.
- 13 R. Illos, F. Bisogno, G. Clodic, G. Bolbach, I. Weissbuch and M. Lahav, *J. Am. Chem. Soc.*, 2008, **130**, 8651.
- 14 R. Illos, G. Clodic, G. Bolbach, I. Weissbuch and M. Lahav, *Origins Life Evol. Biospheres*, 2010, **40**, 51.
- 15 M. M. Green and B. A. Garetz, *Tetrahedron Lett.*, 1984, **25**, 2831.
- 16 M. Bolli, R. Micura and A. Eschenmoser, *Chem. Biol.*, 1997, **4**, 309.
- 17 J. Siegel, *Chirality*, 1998, **10**, 24.
- 18 N. Wagner, B. Rubinov and G. Ashkenasy, *ChemPhysChem*, 2011, **12**, 2771.
- 19 A. Markvoort, H. ten Eikelder, P. Hilbers, T. de Greef and E. Meijer, *Nat. Commun.*, 2011, **2**, 509.
- 20 *Digital Library of Mathematical Functions*, National institute of standards and technology, <http://dlmf.nist.gov/26> technical report.

Models for Mirror Symmetry Breaking via β -Sheet-Controlled Copolymerization: (i) Mass Balance and (ii) Probabilistic Treatment

Celia Blanco and David Hochberg*

Centro de Astrobiología (CSIC-INTA), Carretera Ajalvir Kilómetro 4, 28850 Torrejón de Ardoz, Madrid, Spain

ABSTRACT: Experimental mechanisms that yield the growth of homochiral copolymers over their heterochiral counterparts have been advocated by Lahav and co-workers. These chiral amplification mechanisms proceed through racemic β -sheet-controlled polymerization operative in both surface crystallites as well as in solution. We develop two complementary theoretical models for these template-induced desymmetrization processes leading to multicomponent homochiral copolymers. First, assuming reversible β -sheet formation, the equilibrium between the free monomer pool and the polymer strand within the template is assumed. This yields coupled nonlinear mass balance equations whose solutions are used to calculate enantiomeric excesses and average lengths of the homochiral chains formed. The second approach is a probabilistic treatment based on random polymerization. The occlusion probabilities depend on the polymerization activation energies for each monomer species and are proportional to the concentrations of the monomers in solution in the constant pool approximation. The monomer occlusion probabilities are represented geometrically in terms of unit simplexes from which conditions for maximizing or minimizing the likelihood for mirror symmetry breaking can be determined.



1. INTRODUCTION

Mirror or chiral symmetry is broken in all known biological systems, where processes crucial for life such as replication imply chiral supramolecular structures, sharing the same chiral sign (homochirality). These chiral structures are proteins, composed of amino acids almost exclusively found as the left-handed enantiomers (*S*), and DNA and RNA polymers and sugars with chiral building blocks composed by right-handed (*R*) monosaccharides. The emergence of this biological homochirality in the chemical evolution from prebiotic to living systems is an enticing enigma in the origin of life and early evolution and is a compelling problem that foments scientific activity transcending the traditional boundaries of physics, chemistry, and biology.¹

Biological homochirality of living systems involves large macromolecules; therefore, a key issue is the relationship of the polymerization process with the emergence of chirality. This problem has generated activity in theoretical modeling aimed at understanding mirror symmetry breaking in chiral polymerization. Most of the models proposed^{2,4–10} can be understood as elaborated extensions and generalizations of Frank's original paradigmatic scheme.³ An early work is that of Sandars,² who introduced a detailed polymerization process plus the basic elements of enantiomeric cross inhibition as well as a chiral feedback mechanism in which only the largest polymers formed can enhance the production of the monomers from an achiral substrate. He provided basic numerical studies of symmetry breaking and bifurcation properties of this model for various values of the number of repeat units *N*. The subsequent models cited below are actually variations on Sandars' original theme. Thus, Brandenburg and co-workers⁵ studied the stability and conservation properties of a modified

Sandars' model and introduce a reduced *N* = 2 version including the effects of chiral bias. They included spatial extent⁴ in this model to study the spread and propagation of chiral domains as well as the influence of a background turbulent advection velocity field. The model of Wattis and Coveney⁶ differs from Sandars' in that they allow for polymers to grow to arbitrary lengths *N*, and the chiral polymers of all lengths, from the dimer and upward, act catalytically in the breakdown of the achiral source into chiral monomers. An analytic linear stability analysis of both the racemic and chiral solutions is carried out for the model's large *N* limit, and various kinetic time scales are identified. The role of external white noise on Sandars-type polymerization networks including spatial extent has been explored by Gleiser and co-workers: the *N* = 2 truncated model cited above⁴ is subjected to external white noise,⁷ and chiral bias is then considered,⁸ as well as the influence of high intensity and long duration noise.⁹ Modified Sandars-type models with spatial extent and external noise¹⁰ are considered, allowing for both finite and infinite *N*, with an emphasis paid to the dynamics of chiral symmetry breaking. On the experimental side, Luisi et al.^{11,12} reported the polymerization of racemic tryptophan, leucine, or isoleucine in buffered solutions which yielded libraries of short oligopeptides in the range of six to ten residues, where the isotactic peptides were formed as minor diastereoisomers, in amounts larger than those predicted by a purely random binomial distribution.

One scenario for the transition from prebiotic racemic chemistry to chiral biology suggests that homochiral peptides

Received: June 8, 2012

Revised: October 30, 2012

Published: November 1, 2012

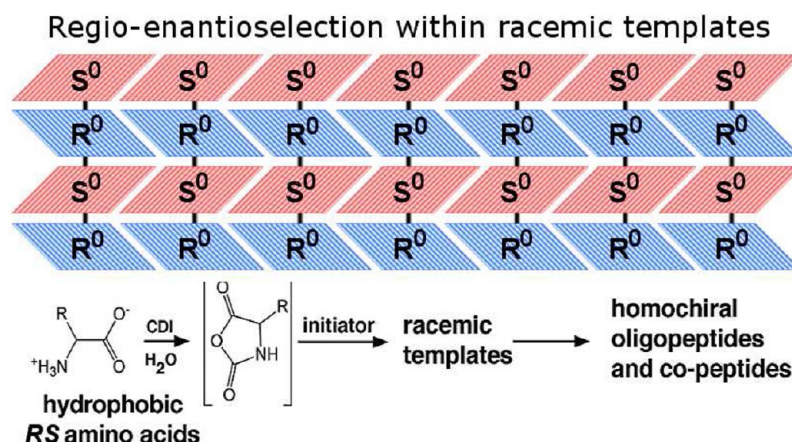


Figure 1. Self-assembly of oligopeptides into racemic β -sheets, for the case of a single species (R_0, S_0) of amino acid supplied in ideally racemic proportions. For a full experimental account, see Weissbuch et al.¹⁹

or amino acid chains must have appeared before the onset of the primeval enzymes.^{13–17} However, except for a couple of known cases,^{11,12} the polymerization of racemic mixtures (i.e., in 1:1 proportions) of monomers in ideal solutions typically yields chains composed of random sequences of both the left and right handed repeat units following a binomial distribution.¹ This *statistical* problem has been overcome recently by the experimental demonstration of the generation of amphiphilic peptides of homochiral sequence, that is, of a single chirality, from racemic compositions or racemates. This route consists of two steps: (1) the formation of racemic parallel or antiparallel β -sheets either in aqueous solution or in 3-D crystals¹⁹ during the polymerization of racemic hydrophobic α -amino acids (Figure 1) followed by (2) an enantioselective controlled polymerization reaction.^{20–26} This process leads to racemic or mirror-symmetric mixtures of isotactic oligopeptides where the chains are composed from amino acid residues of a single handedness (see Figure 1). Furthermore, when racemic mixtures of different types of amino acids were polymerized, isotactic copeptides of homochiral sequence were generated. The guest (*S*) and (*R*) molecules are enantioselectively incorporated into the chains of the (*S*) and (*R*) peptides, respectively; however, the guest molecules are randomly distributed within the corresponding homochiral chains (see Figure 2).

As a combined result of these two steps, the *sequence* of pairs of copeptide *S* and *R* chains within the growing template will differ from each other (see Figure 2). This results in nonracemic mixtures of copeptide polymer chains of different sequences. Consequently, by considering the sequences of the peptide chains, a statistical departure from the racemic composition of the library of the peptide chains is created which varies with chain length and with the relative concentrations of the monomers used in the polymerization.^{21,22} This can be appreciated comparing Figure 1 and Figure 2: in the former the β -sheet is globally racemic (no guest amino acids), whereas the latter template is not by virtue of the randomness of the specific amino acid sequences within each homochiral strand, due to the presence of guest amino acids. It is precisely here, in the β -sheet template, that mirror symmetry is stochastically broken. Non-enantiomeric pairs of homochiral chains are formed; this mechanism relies crucially on the presence of more than one type of amino acid. Note that this does not necessarily imply any net optical activity of solution containing the remaining free chiral monomers.

In this paper, we report a theoretical investigation of multicomponent copolymerization controlled by such templates. The models we introduce presuppose or take as given the prior

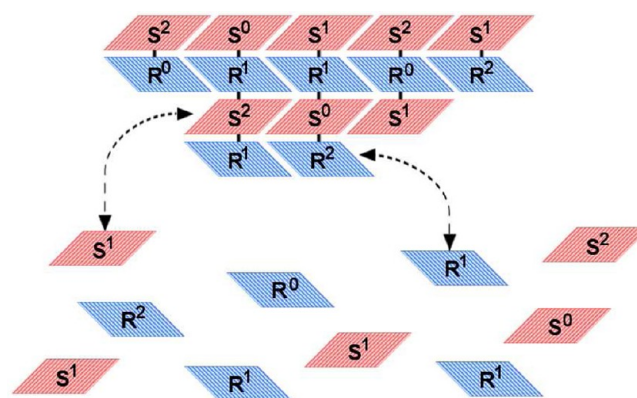


Figure 2. Proposed scheme leading to enantioselective occlusion within racemic β -sheet templates. For the case illustrated, a majority host species (R_0, S_0) and two minority guest species (R_1, S_1) and (R_2, S_2) of amino acids are provided in ideally racemic proportions. The amino acids of a given chirality attach to sites of the same handedness within the growing β -sheet, leading to the polymerization of oligomer strands of a uniform chirality and in the alternating row $S-R-S-R\cdots$ fashion as depicted. Since the polymerization in any given row is random and the guest monomers are typically less abundant than the host species, the former will occlude in a random way leading to independent uncorrelated random sequences in each chiral strand. The overall process yields *nonenantiomeric pairs* of homochiral copolymers, so that mirror symmetry is broken in a stochastic manner. The corresponding mass balance equations (eq 5) are obtained assuming the monomer attachment/dissociation is in chemical equilibrium.

formation of the initial templates or β -sheets and is concerned exclusively with the subsequent enantioselective polymerization reactions. Thus, the nonlinear template control is implicit throughout our discussions. We consider two distinct model approaches to the problem. The first is based on a detailed balance where the polymerization proceeds through stepwise isodesmic additions and dissociations of the chiral monomers (amino acids) to one end of the growing homochiral chain within the template. This can be treated knowing the compositions of the majority and minority monomers and their associated equilibrium constants, and we use chemical kinetics at equilibrium as a useful *approximation* to completely solve the problem. In thermodynamic equilibrium, detailed balance allows us to derive coupled sets of nonlinear mass balance equations. Their solutions yield the equilibrium concentrations of all the monomers and

chiral copolymers in terms of the equilibrium constants and the initial total monomer compositions. With these in hand, we can calculate the enantiomeric excesses of the homochiral chains and their average lengths. The degree of mirror symmetry breaking depends on the numbers of the monomer components and their relative concentrations.²⁷ A brief communication of preliminary results obtained from this equilibrium model were reported recently by the authors.²⁷ This paper extends and generalizes that previous work.

The second model approach is based on strictly probabilistic or statistical considerations and does not assume chemical equilibrium. We will consider the general case involving racemic mixtures of various species of enantiomers, that is, a variety of racemic guest molecules, which can occlude randomly into the chiral sites of the host racemic β -sheet or crystal site.^{21,22,25,26} Among the questions to be addressed: how many species m of such guests are needed to break mirror symmetry? How many repeat units N should the homochiral chains have? What are the ideal mole fractions of the monomers in solution for symmetry breaking? To answer these questions, we first calculate the probability that a given homochiral sequence is formed from the majority and minority species. For random copolymerization, the attachment probability, or the probability of occlusion by the template/crystal, of an amino acid monomer to the growing chain is proportional to its concentration in solution, and we will invoke the constant pool approximation. This probability depends on the polymerization activation energy of the individual monomers (through the Arrhenius relation). The second part deals with combinatorics: counting the number of rearrangements of a given sequence, as all these independent sequences or “reshufflings” will have the same probability to form. The information from both these parts permits us to calculate the joint probability of finding enantiomeric pairs, and from this we deduce the net probability for finding *nonenantiomeric* pairs. The latter provides a statistical measure of the likelihood that mirror symmetry is broken as a function of chain length and the number and concentration of the minority species. We will then generalize these arguments to the case of many additives and even allow nonracemic initial concentrations for all the amino acid species. Here again, the underlying kinetic template control is assumed implicitly.

Both approaches assume that the template-controlled polymerization obeys a first-order Markov process. Experiments carried out in solution^{11,12} appear to confirm this expectation: these results were subsequently rationalized by a mathematical model assuming a first-order Markov mechanism.²⁸

These two theoretical perspectives afford a complementary view of the induced mirror symmetry breaking scenario as originally proposed by Lahav and co-workers. The first scenario holds for closed systems in equilibrium where the monomers are depleted/replenished by the polymerization. We can nevertheless approximate irreversible polymerization as well as we please by simply choosing sufficiently large equilibrium constants. The numerical effects are negligible. The second approach is apt for open systems where the monomer pool is held constant and is free from the assumption of equilibrium.

2. THEORETICAL METHODS I

2.1. Mass Balance. To address the general setting for the generation of libraries of diastereoisomeric mixtures of peptides as originally proposed by Nery et al.,²¹ we need a suitable generalization of their scenario. To this end, we consider the case where we have a majority amino acid species (R_0, S_0) and a given number $m \geq 1$ of minority amino acid species ($(R_1, S_1), (R_2, S_2), \dots, (R_m, S_m)$). Since the following calculations are based on chemical

equilibrium and detailed balance, if all $(m + 1)$ species are supplied in strictly 1:1 racemic proportions, we would justifiably expect a racemic outcome, that is, no mirror symmetry breaking. However, we can test the model's ability for chiral amplification by considering unequal initial proportions for the m minority species in solution. That is, does the enantiomeric excess (ee) increase as a function of chain length, and is it greater than the initial ee of the monomers? The three-monomer case originally treated²¹ is a specific example of this for $m = 1$ and with $R_1 = 0$; that is, the system contains R_0, S_0 and only the enantiomer S_1 of the guest species. We assume as given the prior formation of the initial templates or β -sheets and are concerned exclusively with the subsequent enantioselective random polymerization reactions (step (2)). The underlying nonlinear template control is implicit throughout the discussion. We consider stepwise additions and dissociations of single monomers from one end of the (co)polymer chain, considered as a strand within the β -sheet (see Figure 2). It is reasonable to regard the β -sheet in equilibrium with the free monomer pool.³² (Ref 32 reports a stochastic simulation of two concurrent orthogonal processes: (1) an irreversible condensation of activated amino acids and (2) reversible formation of racemic β -sheets of alternating homochiral strands. The two steps taken together comprise a two-dimensional formulation of the problem. These architectures lead to the formation of chiral peptides whose isotacticity increases with length.)

From a detailed balance, each individual monomer attachment or dissociation reaction is in equilibrium. This holds for closed equilibrium systems in which the free monomers are depleted/replenished by the templated polymerization. Then we can compute the equilibrium concentrations of all the (co)polymers in terms of equilibrium constants K_i for each individual amino acid and the free monomer concentrations. The equilibrium concentration of an S -type copolymer chain of length $n_0 + n_1 + n_2 + \dots + n_m = N$ made up of n_j molecules of type S_j is given by $p_{n_0, n_1, \dots, n_m}^S = (K_0 s_0)^{n_0} (K_1 s_1)^{n_1} \dots (K_m s_m)^{n_m} / K_0$, where $s_j = [S_j]$.²⁹ Similarly for the concentration of an R -type copolymer chain of length $n'_0 + n'_1 + n'_2 + \dots + n'_m = N$ made up of n'_j molecules of type R_j : $p_{n'_0, n'_1, \dots, n'_m}^R = (K_0 r_0)^{n'_0} (K_1 r_1)^{n'_1} \dots (K_m r_m)^{n'_m} / K_0$, where $r_j = [R_j]$. Note that we are considering only copolymers with random sequences such as $R_0-R_0-R_1-R_0-R_0-R_2-R_0-\dots$ and $S_0-S_0-S_1-S_1-S_0-S_2-S_0-\dots$ but not heterochiral polymers (that is, no sequences involving both the S - and R -type monomers). The equilibrium concentration equations we write down, $p_{n_0, n_1, \dots, n_m}^S$ and $p_{n'_0, n'_1, \dots, n'_m}^R$, implicitly assume the underlying template control.

The number of different S -type copolymers of length l with n_j molecules of type S_j for $0 \leq j \leq m$ species, is given by the multinomial coefficient

$$\binom{l}{n_0, n_1, \dots, n_m} = \frac{l!}{n_0! n_1! \dots n_m!} \quad (1)$$

Hence, the total concentration of the S -type copolymers of length l within the β -sheet is given by

$$p_l^S = \sum_{n_0 + n_1 + \dots + n_m = l} \binom{l}{n_0, n_1, \dots, n_m} p_{n_0, n_1, \dots, n_m}^S = \frac{1}{K_0} (K_0 s_0 + K_1 s_1 + \dots + K_m s_m)^l \quad (2)$$

which follows from the multinomial theorem.³⁰ From this we can calculate the number of each type S_j of S -monomer present in the

S-copolymer of length equal to l , for any $0 \leq j \leq m$

$$\begin{aligned} s_j(p_l^S) &= \sum_{n_0+n_1+\dots+n_m=l} \binom{l}{n_0, n_1, \dots, n_m} n_j p_{n_0, n_1, \dots, n_m}^S = s_j \frac{\partial}{\partial s_j} p_l^S \\ &= \frac{K_j}{K_0} s_j (K_0 s_0 + K_1 s_1 + \dots + K_m s_m)^{l-1} \end{aligned} \quad (3)$$

Then, we need to know the total amount of the S-type monomers bound within the S-copolymers (in the β -sheet) from the dimer on up to a maximum chain length N . Using eq 3 for the j th type of amino acid, this is given by

$$s_j(p_{\text{Tot}}^S) = \sum_{l=2}^N s_j(p_l^S) \rightarrow \frac{K_j}{K_0} s_j \frac{a(2-a)}{(1-a)^2} \quad (4)$$

The final expression holds in the limit $N \rightarrow \infty$ provided that $a = (K_0 s_0 + K_1 s_1 + \dots + K_m s_m) < 1$. This must be the case; otherwise, the system would contain an infinite number of molecules.²⁹ Similar considerations hold for the R-sector, and the total amount of R monomers inside R copolymers for the j th amino acid is given by $r_j(p_{\text{Tot}}^R) = (K_j/K_0)r_j(b(2-b)/(1-b)^2)$ where $b = (K_0 r_0 + K_1 r_1 + \dots + K_m r_m) < 1$. From this we obtain the mass balance equations which hold for both enantiomers S,R of the host and guest amino acids and is our key result²⁷

$$s_j + \frac{K_j}{K_0} s_j \frac{a(2-a)}{(1-a)^2} = s_{j_{\text{tot}}}, \quad r_j + \frac{K_j}{K_0} r_j \frac{b(2-b)}{(1-b)^2} = r_{j_{\text{tot}}} \quad (5)$$

These equations express the fact that each type of enantiomer is either free in solution or bound inside a (co)polymer strand within the template.

The problem then consists of the following: given the total concentrations of all the $m+1$ host plus guest enantiomers $\{s_{j_{\text{tot}}}, r_{j_{\text{tot}}}\}_{j=0}^m$ and the equilibrium constants K_j , we calculate the free monomer concentrations in solution $\{s_j, r_j\}_{j=0}^m$ from solving the nonlinear equations (eqs 5). The total system concentration is denoted by $s_{0_{\text{tot}}} + \dots + s_{m_{\text{tot}}} + r_{0_{\text{tot}}} + \dots + r_{m_{\text{tot}}} = c_{\text{tot}}$. From the solutions of eq 5 we can calculate, e.g., the equilibrium concentrations of homochiral copolymers $p_{n_0, n_1, \dots, n_m}^S$ and $p_{n_0, n_1, \dots, n_m}^R$ of any specific sequence or composition as well as the resultant enantiomeric excess for homochiral chains of length l composed of the host (majority) amino acid

$$ee_l = \frac{(r_0)^l - (s_0)^l}{(r_0)^l + (s_0)^l} \quad (6)$$

At this juncture, it is important to point out that our above approach assumes that the polymerization reactions are under thermodynamic control. If there are any kinetic effects, they will not

be seen as they would contribute to the chain compositions at shorter (finite) time scales. Our aim here is to obtain the compositions at asymptotically long relaxation times, and we thus hypothesize that the dominant pathways are under thermodynamic control.

2.2. Average Chain Lengths. We can calculate the average copolymer chain lengths as functions of initial monomer compositions $s_{j_{\text{tot}}}, r_{j_{\text{tot}}}$ and the equilibrium constants K_j , using the solutions of our mass balance equations (eq 5).

The ensemble-averaged chain lengths afford an alternative measure of the degree of mirror symmetry breaking resulting from the desymmetrization process discussed in Nery et al.²¹ There are a number of relevant and interesting averages one can define and calculate. The average chain lengths, starting from the dimers, of the S-type copolymers, composed of random sequences of the S_j type monomers, and that of the R-type copolymers composed of random sequences of the R_j type monomers are given by

$$\begin{aligned} \langle l_S \rangle &= \frac{\sum_{l=2}^N (s_0(p_l^S) + s_1(p_l^S) + \dots + s_m(p_l^S))}{\sum_{l=2}^N p_l^S} \\ &\rightarrow \frac{\left(s_0 + \frac{K_1}{K_0} s_1 + \dots + \frac{K_m}{K_0} s_m\right) \frac{a(2-a)}{(1-a)^2}}{\frac{a^2}{(1-a)K_0}} = \left(\frac{2-a}{1-a}\right) \end{aligned} \quad (7)$$

$$\begin{aligned} \langle l_R \rangle &= \frac{\sum_{l=2}^N (r_0(p_l^R) + r_1(p_l^R) + \dots + r_m(p_l^R))}{\sum_{l=2}^N p_l^R} \\ &\rightarrow \frac{\left(r_0 + \frac{K_1}{K_0} r_1 + \dots + \frac{K_m}{K_0} r_m\right) \frac{b(2-b)}{(1-b)^2}}{\frac{b^2}{(1-b)K_0}} = \left(\frac{2-b}{1-b}\right) \end{aligned} \quad (8)$$

respectively. We also obtain an expression for the average length of the polymer chains composed exclusively from sequences of the S_j or R_j enantiomers of a given specific amino acid of type j

$$\begin{aligned} \langle l_S^j \rangle &= \frac{\sum_{l=2}^N s_j(p_l^{S(s_j)})}{\sum_{l=2}^N p_l^{S(s_j)}} = \frac{\sum_{l=2}^N \frac{K_j}{K_0} s_j (K_j s_j)^{l-1}}{\sum_{l=2}^N \frac{(K_j s_j)^l}{K_0}} \rightarrow \frac{\frac{(s_j K_j)^2 (2 - K_j s_j)}{(1 - K_j s_j)^2}}{\frac{(K_j s_j)^2}{(1 - K_j s_j)}} \\ &= \left(\frac{2 - K_j s_j}{1 - K_j s_j}\right) \end{aligned} \quad (9)$$

$$\begin{aligned} \langle l_R^j \rangle &= \frac{\sum_{l=2}^N r_j(p_l^{R(r_j)})}{\sum_{l=2}^N p_l^{R(r_j)}} = \frac{\sum_{l=2}^N \frac{K_j}{K_0} r_j (K_j r_j)^{l-1}}{\sum_{l=2}^N \frac{(K_j r_j)^l}{K_0}} \rightarrow \frac{\frac{(r_j K_j)^2 (2 - K_j r_j)}{(1 - K_j r_j)^2}}{\frac{(K_j r_j)^2}{(1 - K_j r_j)}} \\ &= \left(\frac{2 - K_j r_j}{1 - K_j r_j}\right) \end{aligned} \quad (10)$$

To complete the list, we can calculate the chain length averaged over all the copolymers in the system

$$\begin{aligned} \langle l \rangle &= \frac{\sum_{l=2}^N (s_0(p_l^S) + s_1(p_l^S) + \dots + s_m(p_l^S) + r_0(p_l^R) + r_1(p_l^R) + \dots + r_m(p_l^R))}{\sum_{l=2}^N (p_l^S + p_l^R)} \\ &\rightarrow \frac{\left(s_0 + \frac{K_1}{K_0} s_1 + \dots + \frac{K_m}{K_0} s_m\right) \frac{a(2-a)}{(1-a)^2} + \left(r_0 + \frac{K_1}{K_0} r_1 + \dots + \frac{K_m}{K_0} r_m\right) \frac{b(2-b)}{(1-b)^2}}{\frac{a^2}{(1-a)K_0} + \frac{b^2}{(1-b)K_0}} \\ &= \frac{a^2(2-a)(1-b)^2 + b^2(2-b)(1-a)^2}{a^2(1-b)^2(1-a) + b^2(1-b)(1-a)^2} \end{aligned} \quad (11)$$

The right-hand most expressions (\rightarrow) in each case hold in the limit of $N \rightarrow \infty$ and for $a < 1$ and $b < 1$. See Table 1 for definitions of all these quantities.

Table 1. Definitions of the Various Average Chain Lengths $\langle l \rangle$, $\langle l_S \rangle$, $\langle l_R \rangle$, $\langle l'_S \rangle$, and $\langle l'_R \rangle$ Employed

$\langle l \rangle$	Average length of all the copolymers in the system
$\langle l_S \rangle$	Average length of all the S-type copolymers, composed of random sequences of the S_j type monomers
$\langle l_R \rangle$	Average length of all the R-type copolymers, composed of random sequences of the R_j type monomers
$\langle l'_S \rangle$	Average length of the polymer exclusively composed from sequences of the S_j enantiomers of a given amino acid type j
$\langle l'_R \rangle$	Average length of the polymer exclusively composed from sequences of the R_j enantiomers of a given amino acid type j

3. RESULTS

3.1. Induced Desymmetrization. We turn to the scenario discussed in Nery et al.²¹ and consider the influence of a single guest species, so $m = 1$ will be sufficient for our purposes. For a single guest, we drop numbered indices and denote the majority host species and concentrations by $r = [R]$, $s = [S]$ and the minority guest with a prime: $s' = [S']$.

We use the above framework to calculate the enantiomeric excess (ee) as a function of chain length l for the three starting compositions of the monomer crystals as reported.²¹ In Figure 3

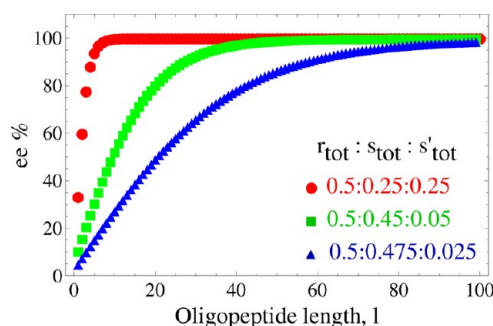


Figure 3. Calculated ee values from solving eqs 5 for the $m = 1$ guest monomer and three different starting monomer compositions (in relative proportions) $r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.25:0.25$ (filled circles), $0.5:0.45:0.05$ (squares), and $0.5:0.475:0.025$ (triangles) for the equilibrium constant $K_0 = K_1 = 1000 \text{ M}^{-1}$ and the total monomer concentration $c_{\text{tot}} = 10^{-3} \text{ M}$. Compare to Figure 13 of Nery et al.²¹

we plot the numerical results obtained from calculating eq 6, and the only quantities required for this are the solutions of r and s obtained from solving the set of equations (eq 5). For strictly illustrative purposes only, we set the equilibrium constants to be the same for both host and guest monomers, $K_1 = K_0 \equiv K = 1000 \text{ M}^{-1}$; the total initial concentration is $c_{\text{tot}} = 10^{-3} \text{ M}$; and the initial fractions of each component are denoted by $f = \{f_r, f_s, f_{s'}\}$ and obey $f_r + f_s + f_{s'} = 1$. The starting composition of the mixture is $c_{\text{tot}} = r_{\text{tot}} + s_{\text{tot}} + s'_{\text{tot}}$ and the total amount of each component is $r_{\text{tot}} = c_{\text{tot}} * f_r$, $s_{\text{tot}} = c_{\text{tot}} * f_s$, and $s'_{\text{tot}} = c_{\text{tot}} * f_{s'}$. We can appreciate the induced symmetry breaking mechanism²¹ from the behavior of ee. For the first case $f_r:f_s:f_{s'} = 0.5:0.25:0.25$, and mirror symmetry is broken for almost all the chain lengths, even for small values of l : for $l = 3$ the ee reaches 60%, and for $l = 5$ the ee is found to be greater than 80%. This is due to the equal starting fractions of the majority s_{tot} and the guest s'_{tot} monomer species of the same chirality, and the large amount of guest is the reason for these

large values of ee. For the second case $f_r:f_s:f_{s'} = 0.5:0.45:0.05$, the starting fraction of the majority species, s'_{tot} is almost 10 times ($0.45/0.05 = 9$) greater than that of the guest, s' . So, for the enantiomeric excess to be greater than 60% the chain length must be at least $l = 13$, and to obtain an ee of 80%, the chain length must be at least $l = 20$. Finally, for the third case, $f_r:f_s:f_{s'} = 0.5:0.475:0.025$, the starting fraction of the majority species, s_{tot} is almost 20 times ($0.475/0.025 = 19$) greater than that of the guest, s'_{tot} , thus the enantiomeric excess for each chain length is expected to be much less than for the two previous cases. An ee greater than 60% is found for the chain length $l = 27$, and for reaching greater than 80%, the chain length must be at least $l = 42$. For the three cases, an increase of the ee is observed (for all l) when increasing the starting fraction of the guest species, s'_{tot} . When s'_{tot} is comparable to s_{tot} while maintaining the proportion R-type:S-type = 1:1, then symmetry breaking is ensured to be >40% for all $l > 5$.

The mass balance equations can be used for calculating the amount of free monomers in solution as well as the amounts of the monomers bound inside the polymers as functions of the starting compositions and K . Solving eq 5 yields the amounts of the free monomers, given by (r, s, s') , while the amounts of the R, S, and S' monomers inside the copolymers are given by the expressions $r_{\text{poly}} = rb(2-b)/(1-b)^2$, $s_{\text{poly}} = sa(2-a)/(1-a)^2$, and $s'_{\text{poly}} = s'a(2-a)/(1-a)^2$, respectively. Then the total amount of all the monomers in polymers is given by $c_{\text{poly}} = r_{\text{poly}} + s_{\text{poly}} + s'_{\text{poly}}$, and the total amount of free monomers in solution is $c_{\text{free}} = r + s + s'$. In Figure 4 we display the values of these quantities for the same three starting compositions considered above as a function of c_{tot} and for $K_0 = K_1 = 1000 \text{ M}^{-1}$. The first row of Figure 4 indicates how the amount of free monomers in solution, c_{free} , is greater than the amount of those in polymers, c_{poly} , for values of c_{tot} below a critical value. Above this value, then $c_{\text{poly}} > c_{\text{free}}$: that is, the majority of the monomers are found in the polymers, not in solution. In the second row, the different contributions to c_{poly} are plotted for each type of monomer. In the first case $s_{\text{tot}}:s'_{\text{tot}} = 1:1$ and leads to $s_{\text{poly}}:s'_{\text{poly}} = 1:1$ which is the most favorable case for mirror symmetry breaking. Increasing the starting ratio between s_{tot} and s'_{tot} increases the difference between s_{poly} and s'_{poly} and diminishes the degree of symmetry breaking. The curves for s_{poly} approach that of r_{poly} as s'_{tot} is diminished (from left to right in Figure 4). Hence, in the third case, where $s_{\text{tot}}:s'_{\text{tot}} = 0.475:0.025$, almost all the monomers present in copolymers are the S monomers. The same applies for the third row, where the different contributions to c_{free} are plotted. Both the amounts of free monomers and those forming polymers increase when increasing c_{tot} . The degree of mirror symmetry breaking can be visualized by the gap or vertical distance between the curves for r_{free} and s_{free} versus c_{tot} and as the amount of s'_{free} is varied. In a similar way, Figure 5 displays the same quantities for fixed $c_{\text{tot}} = 10^{-3} \text{ M}$ and as functions of the equilibrium constant K . As before, c_{poly} and its individual contributions all increase with increasing K , whereas c_{free} (and its individual contributions) all decrease. Clearly, increasing K favors the formation of the polymers over their dissociation into free monomers, and we can approximate irreversible polymerization as close as we please by taking sufficiently large values of K .

The equilibrium concentration of the S-type copolymer chain of length $m + n = N$ made up of m molecules S and of n molecules S' is given by $p_{m,n}^S = (Ks)^m (Ks')^n / K$. Using the solutions from eqs 5 we compute the mole fractions of each S-copolymer, normalized to its own subfamily as $p_{m,n}^S / \sum_{m+n=N} p_{m,n}^S$. This is displayed in Figure 6 for $2 \leq N \leq 7$.

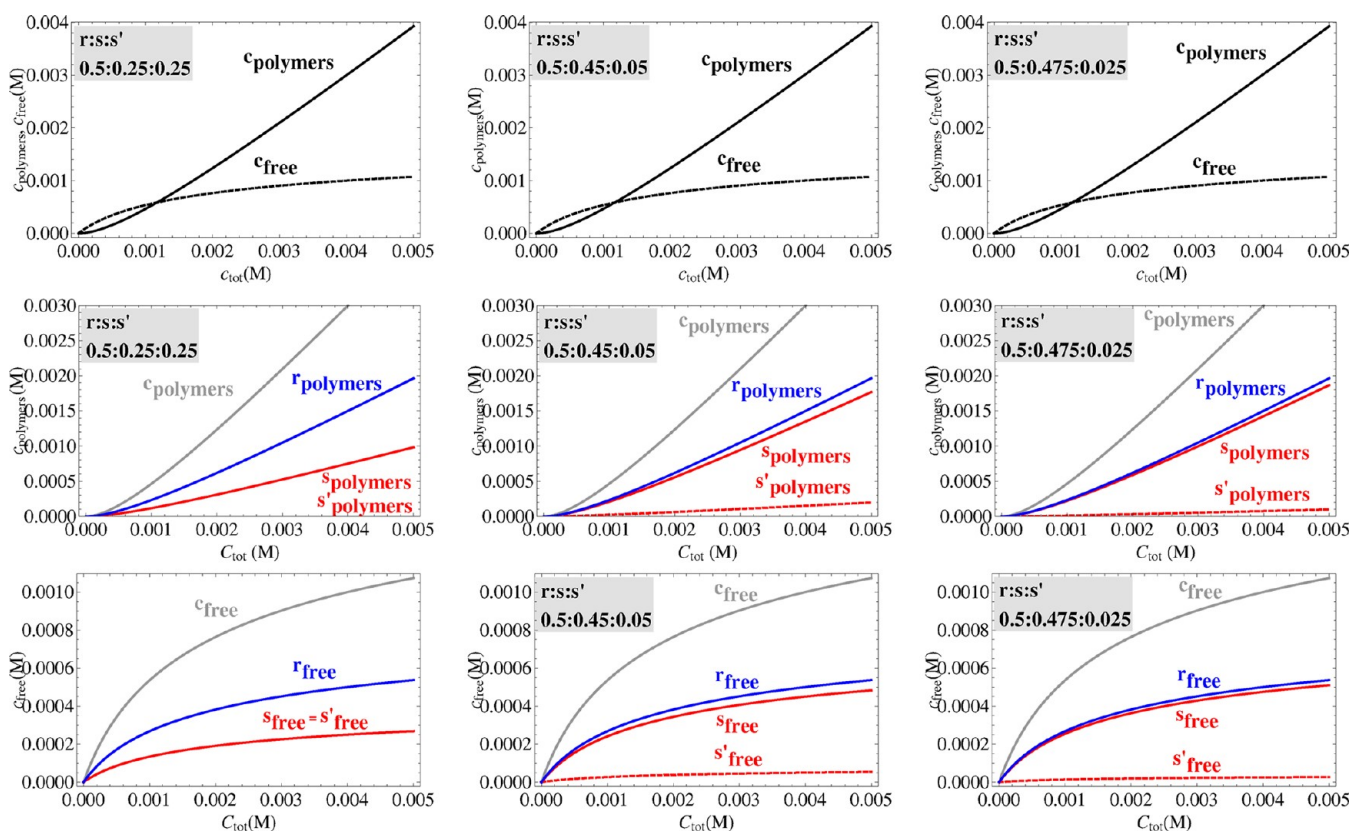


Figure 4. $K_0 = K_1 = 1000 \text{ M}^{-1}$. The amounts of free monomers and those bound in polymers as a function of total monomer concentration c_{tot} versus three different initial relative proportions: $r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.25:0.25$, $r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.45:0.05$, and $r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.475:0.025$. The first row: the total amount of free monomers and those forming polymers. Second row: the total and individual amounts of monomers forming polymers. Third row: the total and individual amounts of free monomers. See text for discussion.

By way of one further example, we carry out a similar analysis for the case of four monomers, this time for two majority R, S and two minority amino acids R', S' . From eq 6 we calculate the e_{ℓ} for the different chain lengths ℓ for three different starting compositions. In Figure 7 we show the numerical results obtained from the solutions of the set of equations (eq 5 and eq 6), for $K_0 = K_1 = 1000 \text{ M}^{-1}$ and $s_{\text{tot}} + s'_{\text{tot}} + r_{\text{tot}} + r'_{\text{tot}} = 10^{-3} \text{ M}$.

As before, we can evaluate the mole fractions of both the S - and R -type copolymers that are in equilibrium with the free monomer pool: namely $p_{m,n}^S / \sum_{m+n=N} p_{m,n}^S$ and $p_{m,n}^R / \sum_{m+n=N} p_{m,n}^R$ respectively. These are displayed in Figure 8 for the initial total compositions indicated there.

Figures 3 and 7 clearly demonstrate that the higher (lower) is the initial degree of chiral asymmetry, characterized by $r_{\text{tot}}/s_{\text{tot}}$ in the former and $r'_{\text{tot}}/s'_{\text{tot}}$ in the latter, and the higher (lower) is the final asymmetry. Thus, rather than symmetry breaking per se, we are observing the model's capacity for asymmetric amplification, as stated at the beginning of Section 2.1. Nevertheless, effects closer to a symmetry breaking effect can be appreciated by looking at the average chain lengths for unequal equilibrium constants in Section 3.2.

3.2. Average Lengths of Copolymer Chains. As an application of the mean chain length formulas derived in eqs 7–11, in the following we focus on the simplest case of the $m = 1$ guest. We consider the effect of different equilibrium constants $K_0 \neq K_1$ and a small total system concentration $c_{\text{tot}} = 10^{-3} \text{ M}$ in Table 2. The dependence on varying c_{tot} for fixed but distinct equilibrium constants $K_0 \neq K_1$ is displayed in Table 3.

Most interestingly, in Tables 2 and 3, one can see the evolution of the global $r/(s + s')$ asymmetry, by looking at the $\langle l_S \rangle / \langle l_R \rangle$

difference. Especially from the results for the 0.5:0.25:0.25 case, i.e. starting from a symmetric $r/(s + s')$ state, some chiral asymmetry, albeit small, is obtained between the length of the all- R and all- S copolymers. The source of this asymmetry is the ratio K_0/K_1 of the equilibrium constants, which we set to 2 in these examples. By contrast, when $K_0 = K_1$ there is then no difference between $\langle l_R \rangle$ and $\langle l_S \rangle$. Conversely, greater ratios of K_0/K_1 lead to greater differences in $\langle l_R \rangle$ and $\langle l_S \rangle$ (data not shown).

Finally Tables 4 and 5 have been calculated for the same starting compositions as Figure 7 and can be compared with the latter.

4. THEORETICAL METHODS II

4.1. Probabilistic Approach. In the following sections, we adopt a statistical approach for calculating the likelihood for finding *nonenantiomeric pairs* of copolymers formed by the proposed template mechanism. This approach does not require chemical equilibrium. How many species m of the chiral guest monomers are needed to break mirror symmetry? How many repeat units N should the chains have? Are there conditions on the polymerization activation energies and mole fractions of the monomers in solution for maximizing the mirror symmetry breaking? We provide answers to these questions based on statistics, and this means being able to count polymer configurations, distinguishing sequences from compositions, and applying some basic combinatorial analysis. Indeed, we may regard the specific homochiral copolymerization sequences formed within the template mechanism as outcomes or “tosses” of generalized multifaceted “dies” (e.g., see Figure 2). However, these dies

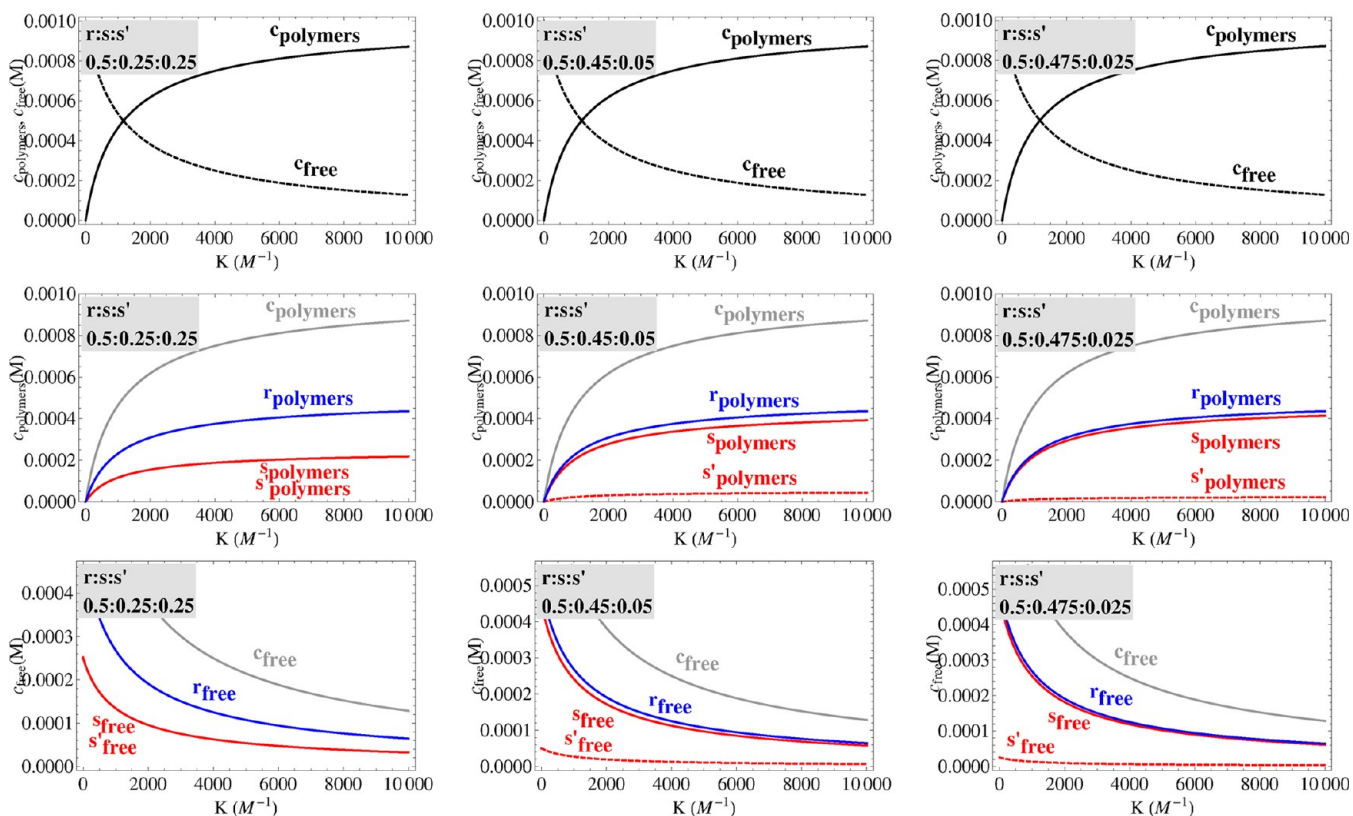


Figure 5. $c_{\text{tot}} = 10^{-3}$ M. The amounts of free monomers and those in polymers as a function of the equilibrium constant K versus three different initial relative proportions: $r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.25:0.25$, $r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.45:0.05$, and $r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.475:0.025$. The first row: the total amount of free monomers and those forming polymers. Second row: the total and individual amounts of monomers forming polymers. Third row: the total and individual amounts of free monomers. See text for discussion.

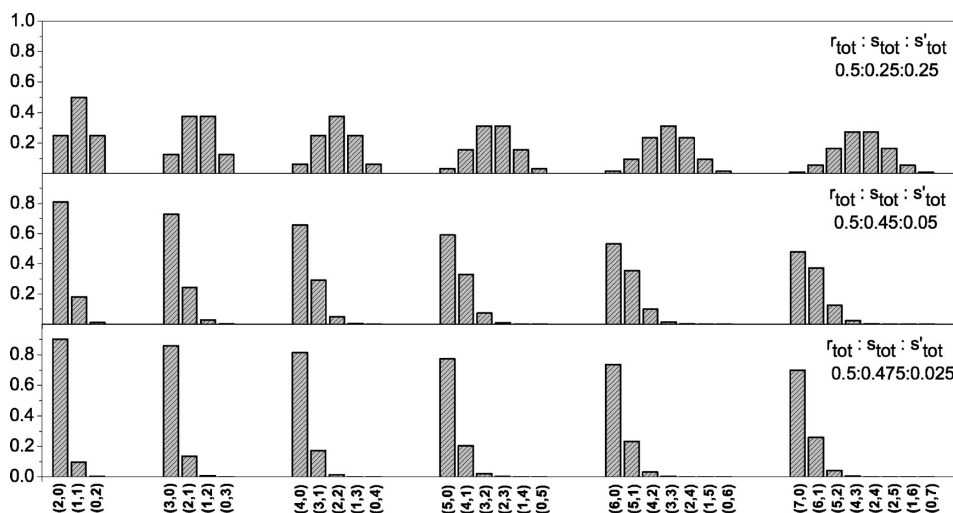


Figure 6. Mole fractions of the S -type copolymers composed of n S and m S' monomers for $m+n=N$ with $2 \leq N \leq 7$, as functions of the total initial fractions $r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}}$. For $s_{\text{tot}}:s'_{\text{tot}} = 1:1$, the distributions are binomial (top), but when $s_{\text{tot}}:s'_{\text{tot}} = 9:1$ (center) or when $s_{\text{tot}}:s'_{\text{tot}} = 19:1$ (bottom), then the distributions are skewed.

are loaded, in the sense that not all faces of the generalized die have an equal probability of turning up in any given throw. This is because different amino acids have different polymerization activation energies and may be present in solution in different proportions.

To resolve this problem, we must pay special attention to both the overall composition of the copolymer chain and its specific sequence. The problem has two basic parts: one is concerned with calculating the probability that a given amino acid sequence

is formed from the majority species and whatever minority species are present in their respective mole fractions in solution. The attachment probability (the probability that the host template occludes this monomer) of an amino acid monomer to the growing chain is proportional to its concentration in solution and to a factor depending on its polymerization activation energy. The second part is to count the number of rearrangements or “shufflings” of the given sequence, as all these

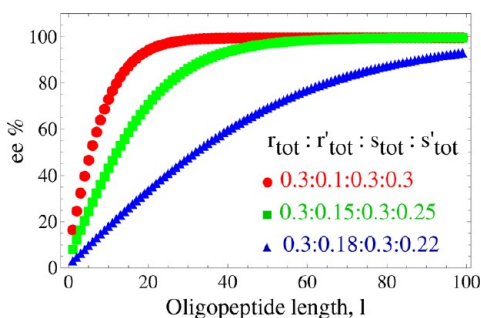


Figure 7. Calculated ee values from solving eqs 5 for three different starting monomer compositions (in relative proportions) $r_{\text{tot}}:r'_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.3:0.1:0.3:0.3$ (filled circles), $0.3:0.15:0.3:0.25$ (squares), and $0.3:0.18:0.3:0.22$ (triangles) for the equilibrium constant $K_0 = K_1 = 1000 \text{ M}^{-1}$ and the total monomer concentration $c_{\text{tot}} = 10^{-3} \text{ M}$.

independent sequences will have the same probability to form as the given one. The information from both these parts will permit us to calculate the joint probability that a given sequence and its mirror image sequence are formed. This in turn will be used to provide a statistical measure of the likelihood that mirror

symmetry is broken: below we derive a compact expression for the probability to find nonenantiomeric pairs of copolymers in the template (Figure 2). We first need to specify the length N of the homochiral copolymer chains to be formed and the number of each minority species or additive m_r, m_s . Thus we consider $(r_0, r_1, r_2, \dots, r_m)$ and $(s_0, s_1, s_2, \dots, s_m)$, whereas $(r, s) \equiv (r_0, s_0)$ denotes both the enantiomers of the majority species. Following the experimental scenario, the minority species will typically be present in small mole fractions, whereas the majority species will be present with a predominantly large mole fraction, as their names suggest. The total number of possible sequences in a chain with N repeat units for each configuration is $(m_r + 1)^N$ and $(m_s + 1)^N$. This most general case is represented in a suggestive pictorial way in Figure 8. This diagram is used to enumerate all possible chiral copolymers that can form in the template, laid out in a linear fashion, the totality of R -copolymers strung out above a “mirror” and the mirror-related S -copolymers directly below it.

Statistical copolymers are those for which the sequence of monomer residues follows a statistical rule. The attachment probability is proportional to the monomer’s concentration in solution $[r_j], [s_j]$ times a rate constant that depends on the

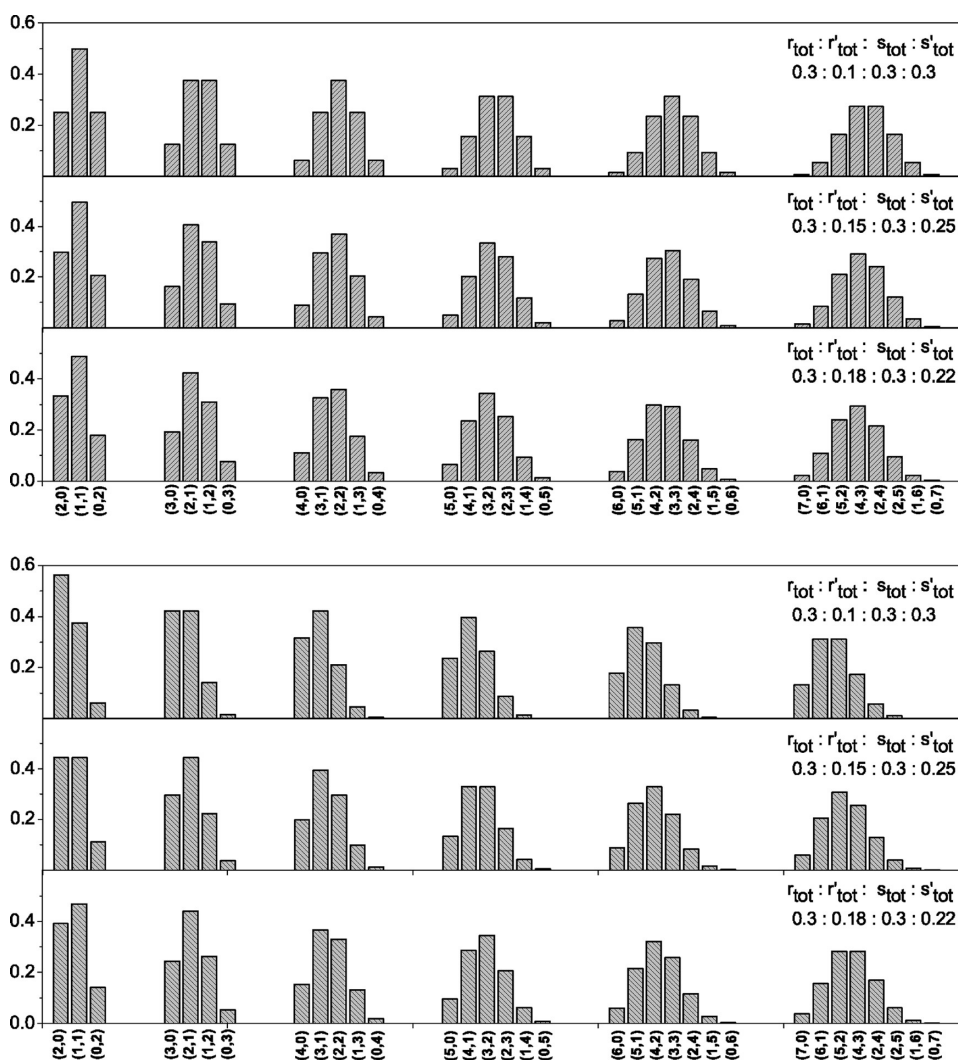


Figure 8. Top set: mole fractions of the S -type copolymers (n, m) for $m + n = N$ with $2 \leq N \leq 7$ (left to right in each row), as functions of the total initial fractions $r_{\text{tot}}:r'_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}}$ indicated there. Bottom set: mole fractions of the R -type copolymers (n, m) for $m + n = N$ with $2 \leq N \leq 7$. Since $s_{\text{tot}}:s'_{\text{tot}} = 1:1$ in the first row of the top graph, the distributions are binomial but are skewed in all the other cases displayed.

Table 2. Average Chain Lengths for the Three Different Starting Compositions As a Function of K_0 for $K_1 = K_0/2$ and $c_{\text{tot}} = 10^{-3}$ M

K_0 (M^{-1})	$r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.25:0.25$					$r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.45:0.05$					$r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.475:0.025$				
	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$
1	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
10	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
100	2.04	2.04	2.05	2.02	2.01	2.05	2.05	2.05	2.04	2.00	2.05	2.05	2.05	2.04	2.00
1000	2.34	2.31	2.37	2.17	2.10	2.36	2.36	2.37	2.32	2.02	2.36	2.36	2.37	2.34	2.00
10000	3.76	3.73	3.79	2.51	2.42	3.79	3.78	3.80	3.40	2.06	3.79	3.79	3.79	3.58	2.03
100000	8.57	8.56	8.59	2.78	2.75	8.59	8.58	8.59	5.60	2.10	8.59	8.59	8.59	6.73	2.04

Table 3. Average Chain Lengths for the Three Different Starting Compositions As a Function of c_{tot} for $K_0 = 100\ 000$ and $K_1 = K_0/2$

c_{tot} (M)	$r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.25:0.25$					$r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.45:0.05$					$r_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.5:0.475:0.025$				
	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$
10^{-5}	2.34	2.31	2.37	2.17	2.10	2.36	2.36	2.36	2.32	2.02	2.36	2.36	2.37	2.34	2.01
10^{-4}	3.76	3.73	3.79	2.51	2.42	3.79	3.78	3.79	3.40	2.06	3.79	3.79	3.79	3.58	2.03
10^{-3}	8.57	8.56	8.59	2.78	2.75	8.59	8.58	8.59	5.60	2.09	8.59	8.59	8.59	6.73	2.04
10^{-2}	23.86	23.86	23.87	2.92	2.91	23.87	23.86	23.87	8.18	2.11	23.87	23.87	23.87	11.93	2.05
10^{-1}	72.21	72.21	72.21	2.97	2.97	72.22	72.22	72.21	9.88	2.11	72.24	72.27	72.21	16.79	2.05

Table 4. Average Chain Lengths for the Two Different Starting Compositions As a Function of K_0 for $K_1 = K_0/2$ and $c_{\text{tot}} = 10^{-3}$ M

K_0 (M^{-1})	$r_{\text{tot}}:r'_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.3:0.1:0.3:0.3$							$r_{\text{tot}}:r'_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.3:0.14:0.3:0.26$							
	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$
1	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
10	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
100	2.04	2.04	2.03	2.03	2.01	2.03	2.00	2.04	2.04	2.04	2.04	2.03	2.01	2.03	2.00
1000	2.33	2.36	2.28	2.19	2.12	2.22	2.04	2.33	2.35	2.30	2.20	2.10	2.22	2.05	2.05
10000	3.77	3.94	3.53	2.53	2.45	2.88	2.16	3.75	3.86	3.61	2.58	2.40	2.78	2.22	2.22
100000	8.62	9.23	7.83	2.80	2.77	3.81	2.27	8.58	8.79	8.13	2.89	2.67	3.44	2.37	2.37

Table 5. Average Chain Lengths for the Two Different Starting Compositions As a Function of c_{tot} for $K_0 = 100\ 000$ and $K_1 = K_0/2$

c_{tot} (M)	$r_{\text{tot}}:r'_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.3:0.1:0.3:0.3$							$r_{\text{tot}}:r'_{\text{tot}}:s_{\text{tot}}:s'_{\text{tot}} = 0.3:0.14:0.3:0.26$							
	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$	$\langle l \rangle$	$\langle l_S \rangle$	$\langle l_R \rangle$	$\langle l'_S \rangle$	$\langle l'_R \rangle$
10^{-5}	2.33	2.36	2.28	2.19	2.12	2.22	2.04	2.33	2.35	2.30	2.20	2.10	2.22	2.05	2.05
10^{-4}	3.77	3.94	3.53	2.53	2.45	2.88	2.16	3.75	3.86	3.61	2.58	2.40	2.78	2.22	2.22
10^{-3}	8.62	9.23	7.83	2.79	2.77	3.81	2.26	8.58	8.97	8.13	2.89	2.67	3.44	2.37	2.37
10^{-2}	23.99	25.99	21.50	2.92	2.92	4.49	2.31	23.90	25.16	22.48	3.06	2.80	3.86	2.43	2.43
10^{-1}	72.58	78.96	64.75	2.97	2.97	4.82	2.33	72.34	76.33	67.83	3.12	2.85	4.05	2.46	2.46

activation energy E_j for attachment of that specific monomer to the polymer/template. Thus

$$p(r_j) \propto A_j \exp(-E_j/kT)[r_j] = w_j[r_j] \quad (12)$$

$$p(s_j) \propto w_j[s_j] \quad (13)$$

To obtain bona fide probabilities, these are normalized so that

$$0 \leq p(r_j) = \frac{w_j[r_j]}{\sum_{k=0}^{m_r} w_k[r_k]} \leq 1, \quad (0 \leq j \leq m_r) \quad (14)$$

$$0 \leq p(s_j) = \frac{w_j[s_j]}{\sum_{k=0}^{m_s} w_k[s_k]} \leq 1, \quad (0 \leq j \leq m_s) \quad (15)$$

which implies

$$\sum_{k=0}^{m_r} p(r_k) = 1, \quad \sum_{k=0}^{m_s} p(s_k) = 1 \quad (16)$$

Normalization ensures that the probability that any single monomer attaches to the template is between zero and unity:

evidently no individual probability can be greater than one, nor can the total probability exceed unity. In writing down eq 12, there are two implicit assumptions being made: (1) the rate of polymerization is independent of polymer length N and (2) the probability of any given monomer joining a polymer is independent of the existing polymer structure. In (1) we are assuming isodesmic polymerization: the successive addition of a monomer to the growing chain leads to a constant decrease in the free energy. This in turn indicates that the affinity of a subunit for a polymer end is independent of the length of the polymer.¹⁸ In (2) we assume the polymerization is a first-order Markov process, and the attachment depends only on the nature of the terminal end of the polymer but not on the monomer sequence in the chain. Evidence for kinetic Markov mechanisms has been observed experimentally in some chiral polymerizations.¹⁹

Define the attachment probability vectors as

$$\vec{p}_R = \{p(r_0), p(r_1), p(r_2), \dots, p(r_{m_r})\} \quad (17)$$

$$\vec{p}_S = \{p(s_0), p(s_1), p(s_2), \dots, p(s_{m_s})\} \quad (18)$$

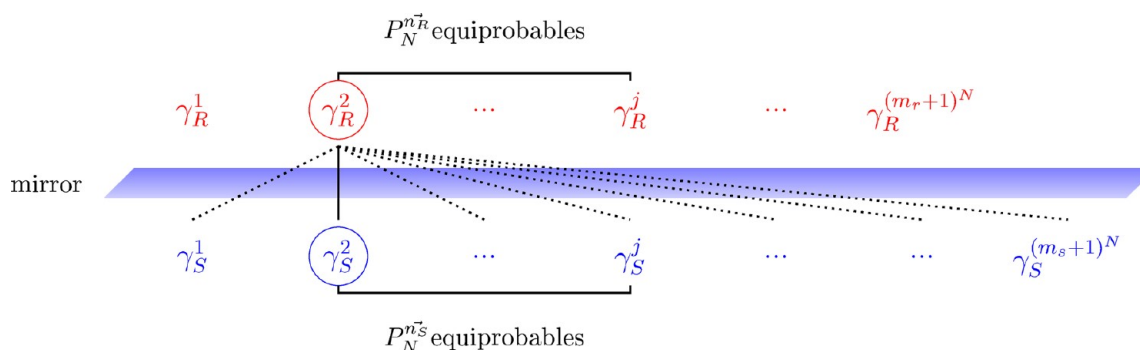


Figure 9. Homochiral copolymer sequences and their mirror-related sequences. Top row (above the mirror) enumerates all the possible R -type copolymers γ_R of length N that can be made up from m_r different R -type monomers: there are $(m_r + 1)^N$ such chains. Below the mirror: mirror image related S -type homochiral copolymers γ_S made up of S -type monomers. In this example, $m_r < m_s$, so there are more S -type copolymers than R -type. The solid vertical line segment links an enantiomeric pair of sequences, and the dotted lines represent examples of *nonenantiomeric* pairs of sequences. A given composition typically gives rise to many inequivalent but equiprobable sequences (indicated by the horizontal solid brackets).

one for the R -monomers and one for the S -monomers. Note that in the limit when both minority species are absent, $m_r, m_s \rightarrow 0$, there will be one unique sequence for each handedness, namely, a sequence of $N R$'s and a mirror sequence of $N S$'s. These two pure sequences will each form with unit probability, since $p(r_0) = p(s_0) = 1$, as follows from eqs 16. An enantiomeric pair will form with absolute certainty when there are no guest additives. This limit provides an important check on the statistical arguments developed below.

Chain compositions for the R - and S -type chains are specified as

$$n(r_0) + n(r_1) + n(r_2) + \dots + n(r_{m_r}) = N \quad (19)$$

and

$$n(s_0) + n(s_1) + n(s_2) + \dots + n(s_{m_s}) = N \quad (20)$$

where $n(r_j)$ and $n(s_j)$ denote the number of times the j th R and S monomer occur in the corresponding chain, respectively. These are ordered partitions of the integer N . Many different *sequences* can follow from one given composition (eqs 19 and 20). By means of the template-controlled polymerization mechanism,²¹ only homochiral chains will be formed, which means that chains formed of either all right-handed R or all left-handed S monomers, and these can be represented by vectors. For example, for the case of a right-handed chain

$$\gamma_R = \{r, r, r_1, r, r, r_3, \dots, r\}_N \quad (21)$$

while its mirror image related *sequence* is denoted by the vector

$$\gamma_S = \{s, s, s_1, s, s, s_3, \dots, s\}_N \quad (22)$$

We emphasize that we are comparing the sequences of copolymers made up exclusively of either all right- R or all left-handed S monomers and not making any claim about their corresponding secondary or tertiary structures. When we discuss copolymers related through the mirror as in Figure 9, we refer exclusively to their specific monomeric sequences or primary structures. It enumerates all the possible sequences that can form from the given composition. The underlying template control is assumed implicitly, thus the system is composed of only homochiral structures (see Figure 2).

The probability to form specific sequences of length N from the compositions 19 and 20 is given by the composition probability

$$p(\gamma_R) = \prod_{j=0}^{m_r} p(r_j)^{n(r_j)} \quad (23)$$

$$p(\gamma_S) = \prod_{j=0}^{m_s} p(s_j)^{n(s_j)} \quad (24)$$

In general, there will be many distinct sequences with exactly the same composition probability (eqs 23 and 24); see the horizontal solid line segments in Figure 9. These are reshufflings or reorderings of the given sequence, keeping the individual composition numbers $n(r_j), n(s_j)$ fixed in eqs 19 and 20, and the number of such equiprobable sequences will be calculated below.

4.1.1. Probability to Form One Enantiomeric Pair. First consider the probability to form a specific sequence, and call it γ_R . We fix the number of repeat units N and the number m_r of R -type additives. From the sequence we immediately deduce the composition (or composition vector) $\vec{n}_R = \{n(r_0), n(r_1), \dots, n(r_{m_r})\}$, and we specify the monomer attachment probabilities (or the attachment/occlusion probability vector) (eq 17). Next, consider the probability to form its mirror image, that is, γ_S . The number of repeat units N has been already fixed, and we know the number m_s of additives of S type. The composition vector of the mirror image \vec{n}_S must be equal to \vec{n}_R that is, $\vec{n}_S = \vec{n}_R \equiv \vec{n}$, and we must specify the monomer attachment probabilities (or the attachment/occlusion probability vector) (eq 18).

The probabilities to form these sequences from these compositions are given by eqs 23 and 24, and hence the joint probability to find the *enantiomeric pair* γ_R and γ_S is

$$P_{\text{pair}}(\gamma_R|\gamma_S) = p(\gamma_R)p(\gamma_S) \quad (25)$$

This is a function of N , $\min(m_r, m_s)$, \vec{n} , \vec{p}_R , and \vec{p}_S .

4.1.2. Probability to Form All Possible Enantiomeric Pairs for Fixed N . For computing the probability of forming all possible enantiomeric pairs, we need to know both m_r and m_s and which one is greater since limits on the possible enantiomeric pairs that can be formed come from the enantiomer with the least number of guest species. Without loss of generality we may assume that $m_r \leq m_s$. For fixed N and m_r , the number of distinct compositions of the R -type copolymers is given by

$$\#_{N, m_r, \{n_0, n_1, n_2, \dots, n_{m_r}\}} = \binom{m_r + N}{N} = \frac{(m_r + N)!}{N!m_r!} \quad (26)$$

and the number of different sequences that we can form from each individual composition is given by

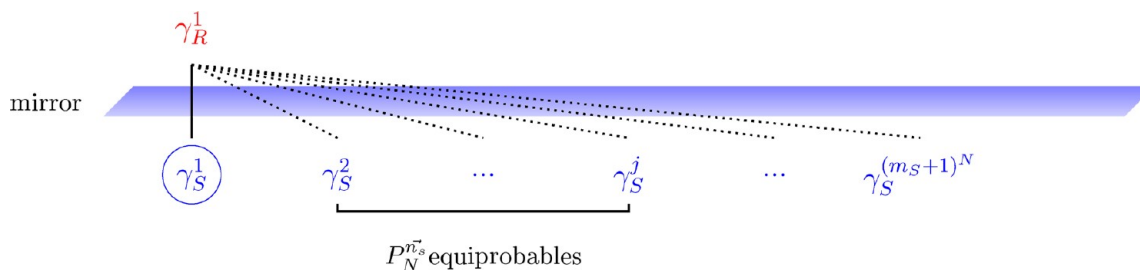


Figure 10. Pictorial situation for chiral guest additives. In this case, only additives of one chirality (in S) are added. So $m_r = 0$ and $m_s = m$, with N fixed. Top row (above the mirror): the unique R -type polymer of length N that can be made up from the single R -type monomer present in the system. Below the mirror: all the S -type copolymers made up of m distinct S -type monomers. Solid vertical line indicates the single unique enantiomeric pair, and the dotted lines represent all the nonenantiomeric pairs of sequences for $m_s \geq 1$.

$$P_N^{\bar{n}_R} \equiv P_N^{n_0 n_1 n_2 \dots n_{m_r}} = \binom{N}{n_0, n_1, n_2, \dots, n_{m_r}} = \frac{N!}{n_0! n_1! n_2! \dots n_{m_r}!} \quad (27)$$

Summing the latter expression over all the possible compositions with fixed N must be equal to the total number of different sequences; that is, we obtain the multinomial theorem³⁰

$$\sum \binom{N}{n_0, n_1, n_2, \dots, n_{m_r}} = (m_r + 1)^N$$

Recall that the joint probability to form a particular sequence and its mirror image sequence is given by eq 25. The net probability we seek to evaluate is

$$P_{\text{pairs}}(N, m_r) = \sum_{\gamma_R} P_{\text{pair}}(\gamma_R | \gamma_S) \quad (28)$$

This expression is the probability that each and every possible sequence in R and its mirror image sequence in S are formed of fixed length N . For this purpose, we will first sum over all different (but equiprobable) sequences belonging to the same composition and then sum over all different compositions for N repeat units. That is $\sum_{\text{all-sequences}} = \sum_{\text{compositions}} (\sum_{\text{equiprobable-sequences}})$. From eq 27 each given composition can be rearranged in $P_N^{\bar{n}_R}$ different ways. For a given composition, all the sequences that can be made therefrom (reshufflings) are equiprobable. Thus, summing over all these possible rearrangements, we arrive at the probability to form chains and their mirror image sequences within one such equiprobable equivalence class. Recall $m_r < m_s$

$$P_N^{\bar{n}_R} p(\gamma_R) p(\gamma_S) = \frac{N!}{n_0! n_1! n_2! \dots n_{m_r}!} \prod_{j=0}^{m_r} p(r_j)^{n(r_j)} \prod_{j=0}^{m_s} p(s_j)^{n(s_j)} \quad (29)$$

Finally, summing this result over all the different compositions, we calculate the net probability to form homochiral chains and their mirror image sequences in the system, i.e., the probability to form all possible enantiomeric pairs. Thus, the probability that mirror symmetry is *not broken* for m additives and N repeat units is given by

$$\begin{aligned} P_{\text{nobreak}}(N, m_r) &= P_{\text{pairs}}(N, m_r) \\ &= \sum_{n_0+n_1+n_2+\dots+n_{m_r}=N} P_N^{\bar{n}_R} p(\gamma_R) p(\gamma_S) \end{aligned} \quad (30)$$

Then the probability that mirror symmetry is broken for these values of m and N is

$$\begin{aligned} P_{\text{break}}(N, m_r) &= P_{\text{nopairs}}(N, m_r) \\ &= \sum_{n_0+n_1+n_2+\dots+n_{m_r}=N} P_N^{\bar{n}_R} p(\gamma_R) (1 - p(\gamma_S)) \\ &= 1 - \sum_{n_0+n_1+n_2+\dots+n_{m_r}=N} P_N^{\bar{n}_R} p(\gamma_R) p(\gamma_S) \\ &= 1 - (p_{r_0} p_{s_0} + p_{r_1} p_{s_1} + \dots + p_{r_{m_r}} p_{s_{m_r}})^N \\ &= 1 - (\bar{p}_R \cdot \bar{p}_S)^N \end{aligned} \quad (31)$$

which follows from the multinomial theorem³⁰ and after using eqs 16, 17, and 18.

4.2. Chiral Additives. This is a special case of the general one described above and is pictorially sketched out in Figure 10. Here we consider $m_r = 0$ and we set $m_s = m$. Clearly, there is only one possible composition (and hence, sequence) that can be formed in R ; namely, the pure homochiral sequence made up of N repeat units of r : $\gamma_R = \{r, r, \dots, r\}_N$, and this forms with unit probability $p(\gamma_R) = 1$. Thus, the probability that mirror symmetry is *not broken* for m types of S -additives and for N repeat units is given by eq 30, which simplifies to give

$$\begin{aligned} P_{\text{nobreak}}(N, m) &= P_{\text{pairs}}(N, m) \\ &= p(\gamma_R) p(\gamma_S) = p(s_0)^N \end{aligned} \quad (32)$$

Then the probability that mirror symmetry is broken for these values of m and N is

$$\begin{aligned} P_{\text{break}}(N, m) &= P_{\text{nopairs}}(N, m) \\ &= p(\gamma_R) (1 - p(\gamma_S)) \\ &= 1 - p(s_0)^N \end{aligned} \quad (33)$$

If the number of S -type additives goes to zero, $m_s \rightarrow 0$, then $p(s_0) \rightarrow 1$ and then mirror symmetry is maintained with absolute certainty.

4.3. Ideal Racemic Additives. For our final example, we deal with the case in which all the additives are supplied in ideally racemic proportions; that is, we have equal numbers of enantiomer types $m_r = m_s \equiv m$, and all are supplied in identical concentrations: $[r_j] = [s_j]$ for all types $0 \leq j \leq m = m_r = m_s$. This situation is pictorially represented in Figure 11. Since the activation energies are the same for both R and S enantiomers, then from eqs 14 and 15 $p(r_j) = p(s_j)$ and from eqs 23 and 24 $p(\gamma_R) = p(\gamma_S) \equiv p(\gamma)$. All the sequences that are obtained from reshuffling the original one have the same probability for

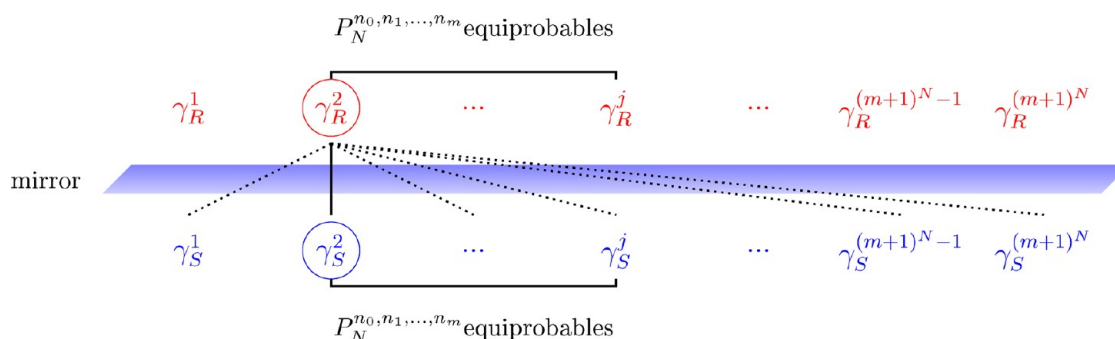


Figure 11. Pictorial scheme for the case of all racemic additives: $m_r = m_s = m$; N is fixed. Compare to Figure 9.

polymerizing: they define equivalence classes of *equiprobable sequences*. This number gives the number of distinct equivalence classes of equiprobable sequences (Figure 11). An important check on the formalism and numerics that follow from it is that the *probability* for mirror symmetry breaking must go to zero as the number m of (ideally racemic) monomer additives goes to zero. We will see that these expectations are confirmed.

For ideally racemic additives, the probability that mirror symmetry is *not broken* for m additives and N repeat units is given by eq 30, which reduces to

$$P_{\text{nobreak}}(N, m) = ((\bar{p}_R)^2)^N \quad (34)$$

Now, from eq 31 the probability that mirror symmetry is broken for m species of racemic additives and for chain length N is

$$P_{\text{break}}(N, m) = 1 - ((\bar{p}_R)^2)^N \quad (35)$$

This result is important: it says that even for ideally racemic initial proportions in all the host and guest amino acids there is a finite probability $1 \geq P_{\text{break}}(N, m) > 0$ for statistical or stochastic breaking of mirror symmetry. This mirror symmetry breaking is manifested in the formation of nonenantiomeric pairs of homochiral sequences within the template, in support of the proposed experimental scenario.²¹

5. RESULTS

From eqs 14 to 18 the monomer attachment probability vectors \vec{p}_R and \vec{p}_S define the faces of two standard or unit m_r and m_s simplexes.³¹ These simplex faces represent the domains of all allowed monomer attachment probabilities (see the shaded regions in Figure 12 and Figure 13). This allows us to find the basic physicochemical criteria for maximizing (or minimizing) the probability for broken mirror symmetry in template-controlled polymerization.^{21,22,25} For an m -simplex there is a maximum and a minimum distance from the origin. The maximum distance pertains when the attachment probability vector \vec{p} coincides with one of the $m + 1$ vertices, and in these cases we have $\|\vec{p}\| = 1$. The minimum distance is achieved for the point defined by the centroid of the simplex face located at $\vec{p} = \{1/(m + 1), 1/(m + 1), \dots, 1/(m + 1)\}$ (a vector with $(m + 1)$ components), and its modulus is $\|\vec{p}\| = 1/(m + 1)^{1/2}$.

5.1. General Case: Nonracemic Additives. For this general case in which the number of additives of each enantiomer type can be distinct, the attachment probability vectors \vec{p}_R and \vec{p}_S have m_r and m_s components, respectively. In principle, they are vectors in simplexes of different dimensions. The limits on the number of possible mirror related copolymer pairs that can be formed come from the enantiomer with the least number of additives, which, without loss of generality, we take to be m_r .

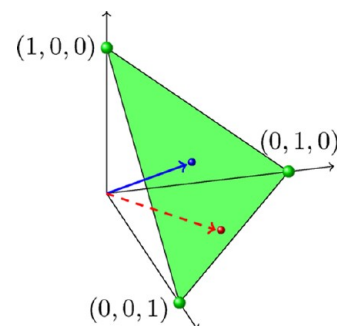


Figure 12. Unit m -simplex, illustrated for the case of $m = 2$ amino acid additives. The three vertices are located at the points $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$ and correspond to maximum attachment probabilities (mirror symmetry is conserved). Shaded area (green) corresponds to the domain of all allowed attachment vectors. A generic point (broken arrow) corresponds to a positive probability for symmetry breaking. The centroid $(1/3, 1/3, 1/3)$ (solid arrow) corresponds to the *maximum* probability for mirror symmetry breaking.

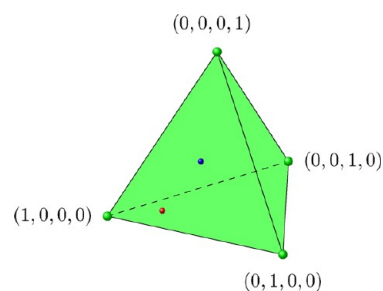


Figure 13. Unit m -simplex, illustrated for the case of $m = 3$ additives. The four vertices are located at the points $(1,0,0,0)$, $(0,1,0,0)$, $(0,0,1,0)$, and $(0,0,0,1)$ and correspond to maximum attachment probabilities (mirror symmetry is conserved). The centroid (blue dot) corresponds to maximal mirror symmetry breaking.

If we express the vector \vec{p}_R in the m_s simplex by simply redefining $\vec{p}_R = \{p_{r_0}, p_{r_1}, \dots, p_{r_{m_s}}, 0, \dots, 0\}$ with $m_s - m_r$ zero entries, then the probability for mirror symmetry breaking follows from eq 31

$$P_{\text{break}}(N, m_r) = 1 - (\vec{p}_R \cdot \vec{p}_S)^N \quad (36)$$

Now, both attachment vectors can be regarded as belonging to the same m_s -simplex. This can be visualized graphically as an m_s -simplex with one of its subspaces being the m_r -simplex (the subspace can be just a point of the m_s -simplex ($m_r = 0$), a line ($m_r = 1$), a face ($m_r = 2$), etc). Figure 14 represents the case in which $m_s = 3$ and $m_r = 2$; here, the subspace of the m_s -simplex

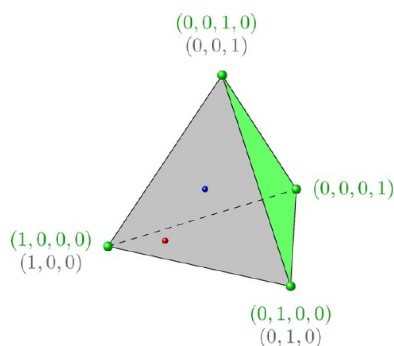


Figure 14. Both the unit m_r - and m_s -simplexes, illustrated for the particular case of $m_r = 2$ and $m_s = 3$ additives. The three vertices of the m_r -simplex are located at the points (1,0,0), (0,1,0), and (0,0,1) and coincide to maximum attachment probabilities of r . These three vertices of the m_r -simplex also coincide with three of the four vertices of the m_s -simplex, located at the points (1,0,0,0), (0,1,0,0), (0,0,1,0), and (0,0,0,1) and corresponding to maximum attachment probabilities of s .

corresponding to the m_r -simplex is a face of the tetrahedron. Then, the probability of breaking symmetry is minimal (zero) when the attachment probability vectors \vec{p}_R and \vec{p}_S are parallel because $\vec{p}_R \cdot \vec{p}_S = 1$ and $P_{\text{break}}(N, m_r) = 0$. For both vectors to be parallel, both must be in the subspace of the m_s -simplex that coincides with the m_r -simplex. The maximum probability for breaking mirror symmetry $P_{\text{break}}(N, m_r) = 1$ is achieved when the attachment probability vectors are orthogonal and thus coinciding with two different vertices of the m_s -simplex (see Figure 13). \vec{p}_R must be in one of the $m_r + 1$ vertices, and \vec{p}_S can be in one of the $m_s + 1$ vertices, always different from the vertex in which \vec{p}_R lies. In this case, the species in r will be different from the species in s , so it will be impossible to form enantiomeric pairs of homochiral chains.

Typically, copolymers will be formed with lengths ranging from the dimer, trimer, etc. on up to a maximum number of repeat units N .^{21,22,25} The above arguments apply to any value of N , thus the probability $P_{\text{break}}^{\leq N}(m)$ to break mirror symmetry in a system containing a spectrum of chain lengths $2 \leq n \leq N$ is given as follows

$$P_{\text{break}}^{\leq N}(m_r) = \frac{1}{N-1} \sum_{n=2}^N P_{\text{break}}(n, m_r) \\ = 1 - \frac{(\vec{p}_R \cdot \vec{p}_S)^2 (1 - (\vec{p}_R \cdot \vec{p}_S)^{N-1})}{(N-1)(1 - \vec{p}_R \cdot \vec{p}_S)} \quad (37)$$

and satisfies $\lim_{\vec{p}_R \vec{p}_S \rightarrow 1} P_{\text{break}}^{\leq N}(m) = 0$ when the two occlusion probability vectors are parallel.

5.2. Additives of Only One Handedness. This is a particular case of the above for $m_r \neq m_s$, when $m_r = 0$. The probability of breaking mirror symmetry depends only on $p(s_0)$: the attachment probability of the S -enantiomer of the majority species. Each monomer attachment vector is in a different simplex with different dimensions: \vec{p}_S is in an m_s -simplex, and \vec{p}_R will coincide with a vertex of the m_s -simplex.

The minimal probability of breaking symmetry $P_{\text{break}}(N, m) = 0$ is obtained for $p(s_0) = 1$; in this case, there are no guests, only the majority species S_0 , so we recover the case in which additives are supplied in racemic proportions. Moreover, no guests are added. In this case, the attachment probability vector \vec{p}_S coincides with one of the $m + 1$ vertices, the vertex corresponding to \vec{p}_R and to $p(s_0)$ maximum. The maximum probability of breaking mirror

symmetry $P_{\text{break}}(N, m) = 1$ is obtained for $p(s_0) = 0$. In this case, the majority species in S is absent, thus no possible enantiomeric pairs can be formed; that is, the vector \vec{p}_S can lie anywhere in the m_s -simplex, except at the m_r vertex.

In this case, the probability $P_{\text{break}}^{\leq N}(m)$ to break mirror symmetry in a system containing a spectrum of chain lengths $2 \leq n \leq N$ (eq 37) reduces to

$$P_{\text{break}}^{\leq N}(m) = 1 - \frac{p(s_0)^2 (1 - p(s_0)^{N-1})}{(N-1)(1 - p(s_0))} \quad (38)$$

and satisfies $\lim_{p(s_0) \rightarrow 1} P_{\text{break}}^{\leq N}(m) = 0$ when no majority species of the S -type is supplied.

The cases with two majority species r and s and one guest, s' , with starting fractions $f_r:f_s:f_{s'}$, as considered in the first section of the paper, would be a case of additives of only one handedness or chiral additive, where $m_r = 0$ and $m_s = m = 1$. Following eq 33 we can calculate $P_{\text{break}}^{\leq N}(m)$ for the three different starting compositions considered before. Exemplary numerical results are shown in Tables 6 and 7

Table 6. $w_r = w_s = w_{s'} = 1$ Probability to Break Mirror Symmetry, $P_{\text{break}}^{\leq N}(m)$, for the Three Different Starting Compositions $f_r:f_s:f_{s'}$ of the Three-Component Case ($m_r = 0$ and $m_s = m = 1$) as a Function of Repeat Units N

$P_{\text{break}}^{\leq N}(m)$	$N = 5$	$N = 10$	$N = 15$	$N = 20$	$N = 25$	$N = 30$
0.5:0.25:0.25	0.88	0.94	0.96	0.97	0.98	0.98
0.5:0.45:0.05	0.30	0.45	0.55	0.63	0.69	0.73
0.5:0.475:0.025	0.16	0.26	0.34	0.41	0.47	0.52

Table 7. $w_r = w_s = 1, w_{s'} = 0.75$ Probability to Break Mirror Symmetry, $P_{\text{break}}^{\leq N}(m)$, for the Three Different Starting Composition $f_r:f_s:f_{s'}$ of the Three-Component Case ($m_r = 0$ and $m_s = m = 1$) as a Function of Repeat Units N

$P_{\text{break}}^{\leq N}(m)$	$N = 5$	$N = 10$	$N = 15$	$N = 20$	$N = 25$	$N = 30$
0.5:0.25:0.25	0.83	0.92	0.95	0.96	0.97	0.97
0.5:0.45:0.05	0.24	0.37	0.47	0.54	0.61	0.66
0.5:0.475:0.025	0.12	0.22	0.27	0.33	0.39	0.43

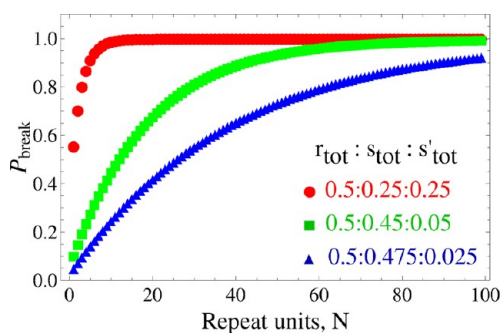


Figure 15. $w_r = w_s = 1, w_{s'} = 0.5$. Probability to break mirror symmetry, $P_{\text{break}}^{\leq N}(m)$, for the three different starting composition $f_r:f_s:f_{s'}$ of the three-component case ($m_r = 0$ and $m_s = m = 1$) as a function of repeat units N .

and in Figure 15, showing the effect of varying the relative concentrations of all the monomers and the activation energy (we vary $w_{s'}$) of the guest monomer s' . The curves for P_{break} are qualitatively similar to those of the percent ee in Figure 3.

5.3. Racemic Additives. When the enantiomeric species are provided in ideally racemic proportions, the probability that

mirror symmetry is broken for given values of the chain length N and number of species m can be expressed succinctly as

$$P_{\text{break}}(N, m) = 1 - \|\vec{p}\|^{2N} \quad (39)$$

That is, one minus the squared-modulus of the probability attachment vector \vec{p} , raised to the chain length. Thus, for fixed N , to maximize the probability that mirror symmetry be broken, we should prepare the chemical system so that all m additives and the majority species have equally shared mole fractions. For any other point in the face (including the centroid), but excluding the $m + 1$ vertices, then $\|\vec{p}\| < 1$, hence the probability to break mirror symmetry increases with chain length N and/or with increasing number of additives m , provided these are supplied with small mole fractions (to prevent \vec{p} from coinciding with the vertices).

Finally, if the occlusion probability vector \vec{p} coincides with any one of the $m + 1$ vertices, then $\|\vec{p}\| = 1$, so $P_{\text{break}}(N, m) = 0$, and mirror symmetry is maintained with absolute certainty for all N . Each vertex corresponds to a chemical system with only one type of monomer (and its enantiomer), in other words, a system with *no additives*. The m th vertex corresponds to the m th amino acid being the sole species present in the system. Thus, we see that if we increase the mole fraction of any one of the additives in excess the tables are turned, and the majority and minority species interchange their roles: excessive amounts of any additive tend to reduce the probability for breaking mirror symmetry.

Equation 37 now simplifies to give

$$P_{\text{break}}^{\leq N}(m) = 1 - \frac{\|\vec{p}\|^4(1 - \|\vec{p}\|^{2(N-1)})}{(N-1)(1 - \|\vec{p}\|^2)} \quad (40)$$

and satisfies $\lim_{\|\vec{p}\| \rightarrow 1} P_{\text{break}}^{\leq N}(m) = 0$ at the vertices of the simplex. As expected, we find increasing probability for symmetry breaking as N and/or m increase. A comparison of the two tables confirms that the probabilities are maximized for each N and m , when all species are supplied in equal proportions. The probability to break mirror symmetry is strictly zero when there are no additives: $P_{\text{break}}^{\leq N}(m = 0) = 0$.

The cases with two majority species r and s and two guests, r' and s' , with starting fractions $f_r, f_{r'}, f_s, f_{s'}$, as considered in the first section of the paper, is a case of racemic additives where $m_r = m_s = 1$. Following eq 35 we can calculate $P_{\text{break}}^{\leq N}(m)$ for the three different starting compositions considered before. The results (not shown) are qualitatively very similar to those shown in the previous tables.

6. CONCLUSIONS

The proposed²¹ experimental mechanism leads to the formation of homochiral copolymers with random sequences of the majority and minority amino acids. Given the implications of the experimental mechanism, we have provided two independent and complementary theoretical approaches to the problem. The first one is based on approximate chemical equilibrium and the second on statistical principles and combinatorics. Both these approaches provide further quantitative insights into the template-controlled induced desymmetrization mechanisms advocated by Lahav and co-workers.^{20–26}

In the first approach, appealing to chemical equilibrium, the template or β sheet is in approximate equilibrium with the free monomer pool. We obtain a multinomial sample space for the distribution of equilibrium concentrations of the homochiral copolymers. We then deduce mass balance equations for the enantiomers of the individual amino acid species, and their numerical solutions are used to evaluate the sequence-dependent copolymer concentrations, in terms of the total species

concentrations. Measurable quantities signaling the degree of mirror symmetry breaking such as the enantiomeric excess (ee), relative abundances, and average chain lengths are evaluated as functions of initial monomer concentrations and the individual equilibrium constants. We can take these constants as large as desired to approximate irreversible polymerization.

The second, or probabilistic, description confirms that this is a viable mechanism for stochastic mirror symmetry breaking. We give criteria for the chemical conditions leading to either maximal or minimal probabilities for breaking mirror symmetry in this experimental context. The probability for finding *nonenantiomeric pairs* of copeptide chains of different sequences increases as a function of increasing chain length and increasing number of guest amino acids. We can calculate the probabilities of all these joint outcomes in terms of the basic monomer attachment/occlusion probabilities. These probabilities can be calculated in terms of the monomer concentrations and take into account the fact that different amino acids have different polymerization activation energies. The solution of the full problem admits an appealing visual and geometric interpretation in terms of unit simplexes which summarize the allowed relative polymerization rates and concentrations of the different amino acids involved.

There are two important points worth emphasizing. First, our theoretical models invoke the underlying template control in that they do not allow for any heterochiral oligomers to form. The sequence of the host and guest amino acids within the homochiral peptides assembles in a completely random fashion, in accord with the experiments.²¹ This sequence randomness is captured by both the model based on chemical equilibrium and the second model based on the monomer occlusion probabilities. Second, the statistical/combinatorial effects do lead to a stochastic mirror symmetry breaking effect. The symmetry breaking in these experiments arises from combinatorics, not from spontaneous (bifurcation) phenomena. These stochastic/statistical/combinatorial effects are not due to the inherent tiny chiral fluctuations present in all real chemical systems^{39,40,43} but are due rather to the random occlusion of host and guest amino acids by the chiral sites of the template: the mechanisms proposed here work even for ideally racemic mixtures. Mirror symmetry is broken in the sequences, as nonenantiomeric pairs of oligomers are formed. The solution of free monomers can nevertheless be optically inactive. The symmetry breaking is to be found in the template, or β -sheet, but not (necessarily) in the solution.

An important distinction must be drawn between the types of symmetry breaking/amplification treated in this paper. Whereas the first part (Section 2) treats the global system symmetry (that thus can lead to global chiral effects), the one described in the latter parts (Section 4) concerns local asymmetries (specific all-*R* versus all-*S* amino acid sequences) that could be “invisible” at the global scale. It is not guaranteed that one asymmetry will imply the other.

The experiments²¹ motivating the present study shed valuable light on the role of templates in the origin of homochiral peptides. Recent works have discussed the potential roles of peptide (amino acid) β -sheets in the origin of life, underscoring their effective protection against decomposition and racemization (recovery of mirror symmetry) as well as their catalytic ability toward hydrolysis.^{33,34} An experimental demonstration of the formation of β -sheets that serve as catalysts for peptide condensation and self-replication has been reported recently.³⁵ Other groups have demonstrated that β -sheet templates can affect enzyme-assisted amino acid polymerization³⁶ and

replication of cyclic peptide structures.³⁷ Such templates might have enjoyed a considerable enantioselective advantage in a prebiotic environment.³⁸

In closing, we note that the symmetry breaking mechanism of Lahav and co-workers^{21,25} has some features in common with the qualitative scenarios of Green, Eschenmoser, and Siegel in which a deficient or limited supply of material results in a stochastic symmetry breaking process.^{41–43}

AUTHOR INFORMATION

Corresponding Author

*E-mail: hochbergd@cab.inta-csic.es.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We are grateful to Meir Lahav (Weizmann Institute of Science) for proposing this problem and for encouraging us to find a mathematical model for the template copolymerization mechanism and for many useful discussions and correspondence. We also acknowledge Gonen Ashkenasy (Ben Gurion University of the Negev) for recent discussions about models for β -sheets. Comments from an anonymous reviewer have helped to improve the manuscript. CB has a Calvo Rodés predoctoral contract from the Instituto Nacional de Técnica Aeroespacial (INTA). DH acknowledges a grant AYA2009-13920-C02-01 from the Ministry of Science and Innovation (currently MINECO).

REFERENCES

- (1) Guijarro, A.; Yus, M. *The Origin of Chirality in the Molecules of Life*; RSC Publishing: Cambridge, 2009.
- (2) Sandars, P. G. H. *Origins Life Evol. Biospheres* **2003**, *33*, 575–587.
- (3) Frank, F. C. *Biochim. Biophys. Acta* **1953**, *11*, 459–463.
- (4) Brandenburg, A.; Multamäki, T. *Int. J. Astrobiol.* **2004**, *3*, 209–219.
- (5) Brandenburg, A.; Andersen, A. C.; Höfner, S.; Nilsson, M. *Origins Life Evol. Biospheres* **2005**, *35*, 225–241.
- (6) Wattis, J. A. D.; Coveney, P. V. *Origins Life Evol. Biospheres* **2005**, *35*, 243–273.
- (7) Gleiser, M.; Thorarinson, J. *Origins Life Evol. Biospheres* **2006**, *36*, 501–505.
- (8) Gleiser, M. *Origins Life Evol. Biospheres* **2007**, *37*, 235–251.
- (9) Gleiser, M.; Walker, S. I. *Origins Life Evol. Biospheres* **2008**, *38*, 293–315.
- (10) Gleiser, M.; Thorarinson, J.; Walker, S. I. *Origins Life Evol. Biospheres* **2008**, *38*, 499–508.
- (11) Blocher, M.; Hitz, T.; Luisi, P. L. *Helv. Chim. Acta* **2001**, *84*, 842–848.
- (12) Hitz, T.; Blocher, M.; Walde, P.; Luisi, P. L. *Macromolecules* **2001**, *34*, 2443–2449.
- (13) Bada, J. L.; Miller, S. L. *Biosystems* **1987**, *20*, 21–26.
- (14) Avetisov, V. A.; Goldanskii, V. I.; Kuzmin, V. V. *Dokl. Akad. Nauk USSR* **1985**, *282*, 115–117.
- (15) Goldanskii, V. I.; Avetisov, V. A.; Kuzmin, V. V. *FEBS Lett.* **1986**, *207*, 181–183.
- (16) Avetisov, V.; Goldanskii, V. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11435–11442.
- (17) Orgel, L. E. *Nature* **1992**, *358*, 203–209.
- (18) De Greef, T. F. A.; Smulders, M. M. J.; Wolffs, M.; Schenning, A. P. H. J.; Sijbesma, R. P.; Meijer, E. W. *Chem. Rev.* **2009**, *109* (11), 5687–5754.
- (19) Weissbuch, I.; Illos, R. A.; Bolbach, G.; Lahav, M. *Acc. Chem. Res.* **2009**, *42*, 1128–1140.
- (20) Zepik, H.; Shavit, E.; Tang, M.; Jensen, T. R.; Kjaer, K.; Bolbach, G.; Leiserowitz, L.; Weissbuch, I.; Lahav, M. *Science* **2002**, *295*, 1266–1269.
- (21) Nery, J. G.; Bolbach, G.; Weissbuch, I.; Lahav, M. *Chem.—Eur. J.* **2005**, *11*, 3039–3048.
- (22) Nery, J. G.; Eliash, R.; Bolbach, G.; Weissbuch, I.; Lahav, M. *Chirality* **2007**, *19*, 612–624.
- (23) Rubinstein, I.; Eliash, R.; Bolbach, G.; Weissbuch, I.; Lahav, M. *Angew. Chem.* **2007**, *119*, 3784. *Angew. Chem., Int. Ed.* **2007**, *46*, 3710–3713.
- (24) Rubinstein, I.; Clodic, G.; Bolbach, G.; Weissbuch, I.; Lahav, M. *Chem.—Eur. J.* **2008**, *14*, 10999–11009.
- (25) Illos, R. A.; Bisogno, F. R.; Clodic, G.; Bolbach, G.; Weissbuch, I.; Lahav, M. *J. Am. Chem. Soc.* **2008**, *130*, 8651–8659.
- (26) Illos, R. A.; Clodic, G.; Bolbach, G.; Weissbuch, I.; Lahav, M. *Orig. Life Evol. Biospheres* **2010**, *40*, 51–63.
- (27) Blanco, C.; Hochberg, D. *Chem. Commun.* **2012**, *48*, 3659–3661.
- (28) Wattis, J. A.; Coveney, P. V. *J. Phys. Chem. B* **2007**, *111*, 9546–9562.
- (29) Markvoort, A. J.; ten Eikelder, H. M. M.; Hilbers, P. A. J.; de Greef, T. F. A.; Meijer, E. W. *Nat. Commun.* **2011**, DOI: 10.1038/ncomms1517.
- (30) Digital Library of Mathematical Functions. National Institute of Standards and Technology. <http://dlmf.nist.gov/26> (accessed Feb 21, 2012).
- (31) Rudin, W. *Principles of Mathematical Analysis*; McGraw-Hill: New York, 1976; Chapter 10.
- (32) Wagner, N.; Rubinov, B.; Ashkenasy, G. *ChemPhysChem* **2011**, *12*, 2771–2780.
- (33) Brack, A. *Chem. Biodiversity* **2007**, *4*, 665–679.
- (34) Maury, C. P. J. *Orig. Life Evol. Biospheres* **2009**, *39*, 141–150.
- (35) Rubinov, B.; Wagner, N.; Rapaport, H.; Ashkenasy, G. *Angew. Chem., Int. Ed.* **2009**, *48*, 6683–6686.
- (36) Williams, R. J.; Smith, A. M.; Collins, R.; Hodson, N.; Das, A. K.; Ulijn, R. V. *Nat. Nanotechnol.* **2009**, *4*, 19–24.
- (37) Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J. P.; Otto, S. *Science* **2010**, *327*, 1502–1506.
- (38) Joyce, G. F.; Visser, G. M.; van Boeckel, C. A.; van Boom, J. H.; Orgel, L. E.; van Westrenen, J. *Nature* **1984**, *310*, 602–604.
- (39) Mills, W. H. J. *Soc. Chem. Ind.* **1932**, *10*, 750–759.
- (40) Mislow, K. *Collect. Czech. Chem. Commun.* **2003**, *68*, 849–864.
- (41) Green, M. M.; Garetz, B. A. *Tetrahedron Lett.* **1984**, *25*, 2831–2834.
- (42) Bolli, M.; Micura, R.; Eschenmoser, A. *Chem. Biol.* **1997**, *4*, 309–320.
- (43) Siegel, J. S. *Chirality* **1998**, *10*, 24–27.

