Review

The Role of Stochastic Models in Interpreting the Origins of Biological Chirality

Gábor Lente

Department of Inorganic and Analytical Chemistry, University of Debrecen, P.O.B. 21, Debrecen, H-4010 Hungary; E-Mail: lenteg@delfin.unideb.hu

Received: 1 March 2010; in revised form: 29 March 2010; / Accepted: 6 April 2010 / Published: 12 April 2010

Abstract: This review summarizes recent stochastic modeling efforts in the theoretical research aimed at interpreting the origins of biological chirality. Stochastic kinetic models, especially those based on the continuous time discrete state approach, have great potential in modeling absolute asymmetric reactions, experimental examples of which have been reported in the past decade. An overview of the relevant mathematical background is given and several examples are presented to show how the significant numerical problems characteristic of the use of stochastic models can be overcome by non-trivial, but elementary algebra. In these stochastic models, a particulate view of matter is used rather than the concentration-based view of traditional chemical kinetics using continuous functions to describe the properties system. This has the advantage of giving adequate description of single-molecule events, which were probably important in the origin of biological chirality. The presented models can interpret and predict the random distribution of enantiomeric excess among repetitive experiments, which is the most striking feature of absolute asymmetric reactions. It is argued that the use of the stochastic kinetic approach should be much more widespread in the relevant literature.

Keywords: absolute asymmetric synthesis; Soai reaction; stochastic kinetics; enantiomeric excess distributions; biological chirality; chiral fluctuations

1. Introduction

Ever since the discovery of molecular chirality, the apparently inherent asymmetric nature of biologically important compounds has attracted the attention of thousands of scientists. In a special issue celebrating the 125th birthday of Science magazine in 2005, its editors mentioned this problem in
a list of 125 most important scientific puzzles that are driving basic experimental and theoretical research [1]. The last comprehensive book on this subject was published in 2004 [2]. It is well known that L-amino acids and D-carbohydrates dominate in life on Earth as we know it. There are very strong reasons to believe that there was no molecular asymmetry in the Universe at some primordial point in time. Therefore, the significant asymmetry of chemical compounds observed today must have been generated at some point by a natural process (probably on Earth), which is thought to be part of evolutionary history. As the very fundamental building blocks of life themselves are chiral, it is just pure common sense to assume that some asymmetry was first generated by pre-biotic chemical processes and later maintained and quite possibly further amplified by biological phenomena. This initial pre-biotic generation of asymmetry has been the subject of extensive theoretical speculation for a long time. Some relevant experimental information reported in the last two decades, especially the discovery of the Soai reaction, has intensified and given new directions to modeling efforts. Asymmetric autocatalysis lies in the center of these efforts, which provides a very attractive interpretation of the eventual origin of biological chirality as, unlike all other proposed explanations, it can be modeled without the inclusion of unknown external factors.

Absolute asymmetric synthesis, in which optically active material is prepared from nonchiral starting materials in the absence of any optically active catalysts or reagents may follow from enantioselective autocatalysis. It should be noted that the requirement for excluding optically active catalysts will be used here in a broad sense that is not limited to chemicals, but also includes circularly polarized light or other radiation, chiral surfaces, or generally any other possible asymmetric external influence that may affect the outcome of a chemical reaction [3]. The traditional way of interpreting absolute asymmetric synthesis relies on the mathematical background of deterministic kinetics and small initial fluctuations, which can never be described based on the deterministic models. The usual course of such interpretations is to show that a particular scheme leads to the formation of huge enantiomeric excesses starting form a minute initial imbalance of enantiomers. Although this is a sound qualitative interpretation, its quantification would require the full description of initial fluctuations, which is not possible within the realms of the deterministic kinetic approach. Consequently, by its very foundations, such a method will be unsuitable for predicting enantiomeric excess distributions observed in experimental examples of absolute asymmetric synthesis. This review will summarize recent stochastic attempts which have been developed to meet the formidable challenges posed by the advance of experimental work in the field.

It should also be noted that the question of how one or other enantiomer built up a reasonable excess from a symmetric initial state during pre-biotic evolutionary processes and reached its present-day levels is quite distinct from the question why exactly the enantiomers in excess today (L-amino acids and D-carbohydrates) were selected by evolution. All reliable information known today points to the direction that this choice between enantiomers was a random one in nature. In other words, life based on D-amino acids and L-carbohydrates would function just as well as the one we are familiar with. Therefore, efforts to answer the first question will be summarized in this article. Only the results of modeling studies will be stated here, the mathematical proofs can be found in the original studies. Mathematical notations will be in agreement with the system used on the web site Wolfram Mathworld [4]. Therefore, equations appearing in the original articles had to be re-formulated from time to time for internal consistency within this review.
2. Stochastic Experimental Observations Relevant for Chirality and Autocatalysis

The reaction most often cited and analyzed in relation to enantioselective autocatalysis is the Soai reaction, which involves a carbon-carbon bond formation reaction using an organozinc compound (Scheme 1).

Scheme 1. The Soai reaction.

\[
\begin{align*}
\text{CHO} & \rightarrow \text{OH} \\
\text{N} \quad \text{N} & \text{and} \\
\text{(C}_2\text{H}_5\text{)}_2\text{O-C}_7\text{H}_8 & \text{trans to the carbaldelyde group on the pyrimidene ring. Later, a two-step synthetic process} \\
& \text{was published in which enormous random fluctuations in the enantiopurity of the product were} \\
& \text{observed and the stochastic nature of the process was demonstrated by 37 repetitive experiments [7].} \\
& \text{The distribution is shown in Figure 1 as series S1. The same reaction was also carried out under} \\
& \text{somewhat different conditions, most notably the presence of nonchiral silica gel was a new element [8].} \\
& \text{With 84 repetitive experiments, this work has yielded the most extensive experimentally determined} \\
& \text{stochastic data set known to date, which is shown in Figure 1 as series S2. Later publications also} \\
& \text{demonstrated that } ^{12}\text{C}/^{13}\text{C} \text{and hydrogen isotope chirality can also be amplified by the Soai} \\
& \text{reaction [9,10]. A review of the experimental results was published within the last year [11].} \\
& \text{Gridnev et al. [12] basically reproduced these findings reported from the Soai group. Their reaction} \\
& \text{was only different from the one shown in Equation 2 by the replacement of the quaternary carbon atom} \\
& \text{by a silicon atom at the non-reactive side chain on the aromatic ring [12]. In that work, 20 repetitive} \\
& \text{experiments gave the distribution shown as series G in Figure 1. The enantiomeric excesses obtained} \\
& \text{are smaller than in S1 and S2, but this could be attributable to the somewhat different substrates and} \\
& \text{the different reaction conditions.} \\
& \text{Singleton and Vo [13] reported the results of 48 synthetic experiments in the same reaction the} \\
& \text{patent from the Soai group studied [6]. These included a lot of thoughtfully devised variation of} \\
& \text{experimental conditions from one experiment to another, which were very useful to make qualitative} \\
& \text{points about the nature of the process but also rendered any statistical attempts to evaluate the data set} \\
& \text{futile. The authors brought up the possibility of a significant role played by minor chiral impurities in} \\
& \text{the system, an important question that, as will be explained in later parts of this review, can actually be} \\
& \text{addressed through statistical analysis. A later paper from the same pair of authors compared 81 separate} \\
& \text{experiments in the same system [14], some of them using the enantioenriched product as a chiral} \\
& \text{additive. The enantiomeric excess values were not specified in this report as the amplification was} \\
& \text{typically repeated until a pre-set excess of one of the two optical isomers were reached.}
Figure 1. Statistical enantiomeric excess distributions ($D$: distribution function) measured in various absolute asymmetric reactions. Origin of data: Reference [7] (S1); Reference [8] (S2); Reference [12] (G); Reference [16] (A).

Mauksch et al. reported random generation of enantiomeric excesses in Mannich and aldol reactions [15] (Scheme 2).

Scheme 2. Absolute asymmetric Mannich and aldol reactions.

This study only demonstrated the existence of a stochastic element in the studied process but did not provide enough repetitive experiments to carry out any meaningful statistical analysis. Even based on the limited data set published [15], there is strong indication that the preference for the $R$ and $S$ enantiomers formed is not 50-50% in this case, which may be attributed to practical imperfections in the experiments rather than an inherent stochastic nature of the system.
In a completely different inorganic reaction, quite substantial random fluctuations were reported in the cleavage of a Co(III)(μ-OH)₂Co(II)(μ-OH)₂Co(III) type trinuclear cobalt complex [16] (Scheme 3).

Scheme 3. Absolute asymmetric cleavage of a trinuclear cobalt complex.

The product octahedral cobalt(III) complex is chiral due to the presence of the ethylenediamine ligands and its enantiomers can be resolved because of its inertness. Small, but clearly detectable random fluctuation was observed in the enantiomeric excess of the chiral product in a series of experiments which involved 20 repetitions of identical procedures. The experimental distribution of determined enantiomeric excess obtained is displayed in Figure 1 as series A. Some earlier and later efforts were also published connected to this particular reaction [17–19].

A major shortcoming of the experimental data reported about absolute asymmetric synthesis is that they were rather focused on the synthetic aspects and the final formation of large random enantiomeric excesses without giving much attention to meaningful kinetic monitoring of the reactions. This is important because the stochastic aspects must be observed in the kinetics of these processes as well. However, experimental examples of detectable stochastic kinetic effects in autocatalytic reactions not involving any chiral material were provided by the pioneering work of Nagypál and Epstein [20–21]. The first such process studied was the reaction between chlorite ion and thiosulfate ion, the stoichiometry of which can be described by the combination of the following two reactions depending on the experimental conditions [20]:

$$4\text{S}_2\text{O}_3^{2-} + \text{ClO}_2^- + 2\text{H}_2\text{O} = 2\text{S}_4\text{O}_6^{2-} + 4\text{OH}^- + \text{Cl}^-$$

(5)

$$\text{S}_2\text{O}_3^{2-} + 2\text{ClO}_2^- + \text{H}_2\text{O} = 2\text{SO}_4^{2-} + 2\text{Cl}^- + 2\text{H}^+$$

(6)

In unbuffered solution with a pH around 11, the reaction shows Landolt or “clock reaction” characteristics in a rigorous sense [22], whereby an initial increase in pH is followed by an abrupt drop. The Landolt time was shown to display stochastic fluctuations and a statistically meaningful study (at least 50 repetitions of identical experiments for characterizing each distribution) at various experimental conditions, including initial concentrations, stirring rates, reactant volumes and medium temperatures was carried out in a systematic manner. The distribution of Landolt times from one of the experimental series is shown by curve N1 in Figure 2. A similar full-scale stochastic distribution study was also reported in the reaction between chlorite ion and iodide ion [21], where the gradual buildup in iodine is followed by its sudden disappearance in an acetate-buffered solution at pH ~ 4 because of the coupling of the following two processes:

$$4\text{I}^- + \text{ClO}_2^- + 4\text{H}^+ = 2\text{I}_2 + \text{Cl}^- + 2\text{H}_2\text{O}$$

(7)
\[ I_2 + 5\text{ClO}_2^- + 2\text{H}_2\text{O} = 4\text{IO}_3^- + 5\text{Cl}^- + 4\text{H}^+ \]  

Figure 2. Statistical reaction time distributions (D: distribution function) measured in the chlorite ion–thiosulfate ion (N1) [20] and chlorite ion–iodide ion (N2) [21] reaction.

There are numerous examples of chiral symmetry breaking during crystallization, which are somewhat similar to the homogeneous examples listed above. It should be made clear that a similar phenomenon, spontaneous resolution in crystallization is quite different conceptually. In this latter process, which is the historical way of the discovery of chirality [23–24], an existing racemic mixture (such as tartaric acid) is crystallized into enantiopure crystals and there is no need for any random elements. In chiral symmetry breaking crystallization, the source of chirality is the formation of enantiomorphic crystals (similar to quartz) and the initial substance may not be chiral at all. The stochastic nature was nicely demonstrated in the crystallization of sodium chlorate [25] and amino acids glutamic acid and lysine [26]. Relevant results have been reviewed from several different aspects [27–32] and extended to phase transitions in general [33–34].

3. Statistical Tests without Chemical Models

A limited analysis of the experimentally measured statistical distributions is possible without having any particular chemical mechanism in mind. As there is no limit on the number of repetitive experiments one can do, comparisons are routinely done with continuous statistical distributions. Data set A displayed in Figure 1 has been shown to adhere to normal distribution [35]. The probability and distribution functions are as follows:

\[ P(x_R) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x_R-\mu)^2}{2\sigma^2}} \]  

\[ D(x_R) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{x_R} e^{-\frac{(\tau-\mu)^2}{2\sigma^2}} d\tau \]

The distributions are given here in terms of the molar fraction of the R isomer (x_R), which is connected to the percentage enantiomeric excess (ee%) with the following simple formula:
If \( x_R < 0.5 \), the \( S \) enantiomer is in excess, conversely, the \( R \) enantiomer is in excess for \( x_R > 0.5 \). Distribution A shown in Figure 1, which is in a narrow range around the enantiomeric excess of zero, was assumed to be centered around the racemic mixture. etc. the expectation \( \mu = 0.5 \) was pre-set and a standard deviation of \( \sigma = 0.037 \) was found to fit the experimental data best. [35] (Note that the value of \( \sigma \) was reported erroneously by a multiplication factor of \( \sqrt{2} \) in the original article [35].) The theoretical distribution is shown as a solid line in Figure 1. The normal distribution approximates the experimental measurements in their range well, but the adherence should not be taken literally outside the range of experimental data, as Equation 9 and Equation 10 would give nonzero probabilities (2.1 \( \times \) 10\(^{-42} \) in this case) for the impossible case where the enantiomeric excess is larger than 100%. Later distributions appearing in this review are free from this problem because they are only interpreted on a limited interval of real numbers. The good fit to the normal distribution in the case of data set A is not unexpected given the relatively minor fluctuations measured.

A combined beta distribution has been proposed to interpret the experimental distributions found in the Soai reaction with the following probability and distribution functions [36]:

\[
P(x_R) = w \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} x_R^{b-1}(1-x_R)^{a-1} + (1-w) \frac{\Gamma(c + d)}{\Gamma(c)\Gamma(d)} x_R^{d-1}(1-x_R)^{c-1}
\]

(12)

\[
D(x_R) = w \frac{B(x_R,a,b)}{B(a,b)} + (1-w) \frac{B(x_R,c,d)}{B(c,d)}
\]

(13)

In these formulae, \( \Gamma(a) \) is the gamma function, \( B(x_R,a,b) \) is the incomplete beta function and \( B(a,b) \) is the complete beta function. As shown in Figure 1, a good-looking fit to data set S2 could be obtained with the values \( a = 2, b = 10, c = 5, d = 1 \) and \( w = 0.382 \). (This is slightly different order from the one reported in the original work [36] as the left part of Figure 1 refers to cases where the \( S \) isomer is in excess.) The probability function based on Equation 11 was compared [36] with a histogram constructed based on the published data [8]. This is an unfortunate practice as it includes some rather arbitrary categorization, the dangers of which have already been pointed out [35]. To demonstrate the goodness of fit, a comparison between distribution functions seems to be the best practice thus far [20,21,35,37,38]. Furthermore, it should be noted that distribution (12) is clearly asymmetric, the significance of which will be discussed later along with distributions giving a visually more attractive fit to data set S2 with fewer parameters, both symmetric and asymmetric, based on chemical models [35,37,38].

The symmetry of the experimentally measured distributions can be tested independently of any chemical considerations [35,38,39–41]. The Wilcoxon rank sum test showed that there is no significant asymmetry in data sets A and S1 [35]. Data set S2, however, was shown to be significantly asymmetric by both the simple Wilcoxon rank sum test [38] and by more advanced statistical evaluation [39–41]. Furthermore, it was shown that (a) distributions S1 and S2 are non-binomial (hence not normal), (b) S2 cannot be described by a symmetric beta distribution [39]. During the present general analysis, it should be pointed out that property (b) also follows from the fact that distribution S2 is significantly asymmetric.
4. Model-Based Stochastic Kinetic Approaches

In many fields of science, it seems evident that stochastic experimental phenomena must be interpreted by stochastic models. Unfortunately, this is not the case in the quest for the origins of biological chirality where published papers based on the deterministic approach outnumber stochastic attempts by at least two orders of magnitude. The reason for this rather puzzling insufficiency might be that the problem is distinctly chemical, but chemists, unlike physicists and biologists, are usually trained all over the world without much exposure to basic probability theory. It should be made clear that the stochastic approach to chemical kinetics is already well established. As early as 1940, Delbrück noted that inherent stochastic phenomena are expected under certain conditions in autocatalytic reactions dealing with macroscopic amounts of substance [42], an insight which was experimentally verified in work already discussed in Section 2 [20,21]. Stochastic approaches to chemical kinetics were nicely developed and summarized by Érdi and Tóth [43,44]. Stochastic kinetics is routinely used in physics to evaluate measurements studying rare radioactive decays [45] and stochastic considerations are also routine in the physical analysis of symmetry breaking and restoration [46].

First of all, a difference between irreproducibility and inherent stochastic nature should be clearly made in chemical systems [35,47]. Both phenomena imply random variations in experimental results, but the underlying reasons are very different. Irreproducibility is caused by insufficiently controlled external factors, e.g. impurities, environmental temperature or pressure fluctuations, light absorption, etc. In such cases, recognizing the important external factor and improving its control will result in greatly reduced random variations. Inherent stochastic nature, e.g. a movement of an electron on a quantum mechanical level, is due to the laws of nature and no refinement of experimental control will remove it. However, the laws governing the behavior of the system must be reflected in experimental results. In relation to enantioselective autocatalysis, four criteria were proposed to differentiate between irreproducibility and stochastic nature [35]:

1. Traditional demonstration of autocatalysis by experiments using the initial addition of enantiomer-enriched products. Autocatalysis is a possible reason for inherent stochastic nature in macroscopic experiments, but autocatalysis can be tested under circumstances where no stochastic behavior is expected.

2. Analysis of the stochastic distribution of enantiomers through a high number (preferably > 50) of repetitive experiments. In most cases, symmetric distributions are expected and a significantly asymmetric experimental distribution may point to an uncontrolled (most often also unidentified) but significant external influence.

3. Detection of random fluctuations in reaction time. If a random enantiomeric distribution is indeed caused by inherent stochastic nature, the time-dependent kinetics should also show random variations.

4. The final distribution of enantiomeric excess and reaction times should be dependent on the overall volume even if the very same initial concentrations are used. This is a theoretically well known phenomenon that is usually expressed by saying that stochastic kinetics is equivalent to deterministic kinetics at the limit of infinite volume [48].
It should be noted that the eventual inevitability of the use of stochastic kinetics for certain problems can be forcefully demonstrated by pointing out the insufficiencies of the deterministic approach. Consider the following pair of simultaneous differential equations, which give the deterministic description for simple first-order enantioselective autocatalysis transforming a nonchiral A molecule into BR and BS enantiomers of a chiral B product [37]:

\[
\frac{d[B_R]}{dt} = 0.5k_u ([A]_0 - [B_R] - [B_S]) + k_c ([A]_0 - [B_R] - [B_S]) [B_R] \quad (14)
\]

\[
\frac{d[B_S]}{dt} = 0.5k_u ([A]_0 - [B_R] - [B_S]) + k_c ([A]_0 - [B_R] - [B_S]) [B_S] \quad (15)
\]

In these equations, \(k_u\) and \(k_c\) represent rate constants, \([A]_0\) is the initial concentration of reactant A, mass conservation is assumed to hold for the sum of substances A and B, \([B_R]\) and \([B_S]\) give the concentrations of the two enantiomers as a function of time. These equations can be solved analytically [37]. Parameters \([A]_0 = 0.2 \text{ M}, k_u = 10^{-15} \text{ s}^{-1}\) and \(k_c = 10^8 \text{ M}^{-1} \text{s}^{-1}\) will be used in the example. The rate constants have been chosen so that the deterministic approach leads to an irresolvable contradiction. However, the values assumed for \(k_u\) and \(k_c\) are by no means impossible from a practical point of view. The value of \(k_u\) corresponds to an extremely slow first-order process, whereas the other path represents a fast second-order reaction with a \(k_c\) that is about two orders of magnitude lower than the diffusion-controlled limit. If there is absolutely no chiral material present initially, etc. \([B_R](t = 0) = [B_S](t = 0) = 0\), the symmetry of Equation 14 and Equation 15 ensures that the final state will also be symmetric with \([B_R](t \to \infty) = [B_S](t \to \infty) = 0.1 \text{ M}\). However, the very same parameters with the initial condition representing a single BR molecule in 10 cm\(^3\) of overall reaction solution, \([B_R](t = 0) = 1.66 \times 10^{-22} \text{ M}\) and \([B_S](t = 0) = 0\) will predict \([B_R](t \to \infty) = 0.1943 \text{ M}\) and \([B_S](t \to \infty) = 0.0057 \text{ M}\). This is a huge contradiction as the first molecule formed in the system must be either BR or BS, there is no middle ground there, so the system must go through either the asymmetric state mentioned in the previous analysis or its mirror-image counterpart \([B_R](t = 0) = 0\) and \([B_S](t = 0) = 1.66 \times 10^{-22} \text{ M}\). Therefore, the final enantiomeric balance predicted based on the initial state and on an intermediate state that the system must necessarily go through contradict each other.

An apparent escape of this contradiction might be thinking that it is always imperative for two molecules of B (one BR and one BS) to form every time at exactly the same moment, but this would be tantamount to denying the very foundations of atomistic theory (independency of individual particles). In probabilistic terms, two molecules formed at exactly the same time is the surprising, but statistically well known case of an event that has zero probability but not impossible.

This example clearly demonstrates that the major insufficiency of the deterministic kinetic approach is its reliance on continuous concentration-time functions. As the usual number of particles dealt with in chemistry is very high (about \(10^{20}\) or higher), this approach is quite satisfactory for the overwhelming majority of cases. However, one should be vigilant about cases where it cannot work exactly because of the initial (and often tacit) assumptions. As the above example shows, it would be a mistake to think that the stochastic kinetic effects arising from the particulate nature of matter can only be significant at unpracticably low molecule numbers, autocatalysis and its consequences clearly prove the opposite.
There are several different methods to account for stochastic elements in chemical kinetics. The most important is called the continuous time discrete state (CDS) approach and will be discussed in the remaining sections of this review. In this section, however, a brief summary of different attempts will be provided.

It has been pointed out several times that a racemic mixture in dynamic equilibrium does not contain strictly the same number of both enantiomer molecules \([3,40,49–59]\). A racemic mixture in general can be characterized by a binomial distribution. This could be well understood within the concepts of the CDS approach but will be discussed in this Section as it has no kinetic relevance at all. The binomial distribution for the case where some inherent asymmetry between enantiomers is also allowed takes the following form \([54]\):

\[
P(r,s) = \binom{r+s}{r} (0.5 + \varepsilon)^r (0.5 - \varepsilon)^s
\]

\(P(r,s)\) is the probability of obtaining exactly \(r\) molecules of the R (more stable) enantiomer and \(s\) molecules of the S (less stable) enantiomer. Parameter \(\varepsilon\) is characteristic of the degree of inherent asymmetry \((\varepsilon \leq 0.5)\). Any formula given in the following equations can be used for the perfectly symmetric case with setting \(\varepsilon = 0\). The importance of calculations in the asymmetric case was that the possible consequences of the parity violation energy difference had to be analyzed. Parameter \(\varepsilon\) can be calculated from the inherent energy difference \((\Delta E)\) between the enantiomers \([54]\), the temperature \((T)\) and the universal gas constant \((R)\) as follows:

\[
\varepsilon = \frac{\Delta E}{4RT}
\]

For a given overall number of chiral molecules \(N (= r + s)\), the expectations and standard deviation of the of the distribution given in Equation 16 can be given as:

\[
\langle r \rangle = (0.5 + \varepsilon)N \\
\langle x_R \rangle = 0.5 + \varepsilon \\
\sigma_r = \sqrt{N(0.5 + \varepsilon)(0.5 - \varepsilon)} \\
\sigma_{x_R} = \frac{(0.5 + \varepsilon)(0.5 - \varepsilon)}{N}
\]

Present theoretical predictions \([60]\), which have been especially carefully checked for the product of the Soai reaction \([61]\), show that \(\Delta E\) following from the phenomenon of parity violation should be around \(10^{-13}\) J mol\(^{-1}\) giving an estimate of \(10^{-17}\) for \(\varepsilon\) at room temperature. It has been pointed out several times \([54,58,59]\) that for 1 mol of chiral molecules \((N = 6 \times 10^{23})\) these values lead to an expected excess of the more stable (R in this case) enantiomer \((6 \times 10^6)\) that is much smaller than the standard deviation \((1.9 \times 10^{11})\). The present author believes that any role played by the parity violation energy differences is highly unlikely based on this simple line of thought alone.

As a further stochastic analysis, the probability of forming the more stable enantiomer in excess \((W)\) was defined and calculated analytically \([54]\). Odd and even numbers of \(N\) had to be treated separately, which is understandable as for an odd number of molecules it is impossible to have a perfectly racemic mixture \((r = s)\) by default. With a unified treatment, a new parameter \(l\) is introduced which is given as follows:

\[
l = \begin{cases} 
N/2 - 1 & \text{for even } N \\
(N - 1)/2 & \text{for odd } N
\end{cases}
\]
The probability \( W \) can be given as [47]:

\[
W = 0.5 + 0.5 \times \sum_{i=0}^{l} \left( P(N - r, r) - P(r, N - r) \right) = \\
= 0.5 + 0.5 \times \sum_{i=0}^{l} (-1)^i e^{2i+1} \frac{(2i+1)!}{(2i+1)! (i!)^2} \approx 0.5 + 2e \sqrt{\frac{N}{2\pi}}
\]

(21)

The last part of Equation 21 is an approximation that only works well for \( W \) values smaller than 0.6. Probabilities \( W \) (actually \( W - 0.5 \) to have a practical logarithmic scale) are shown as a function of \( N \) in Figure 3. Further characteristics of the distribution have been estimated. The probability of forming a perfectly racemic \((r = s)\) mixture can be found from Equation 16 by substituting \( r = s \) [46]. Mills [43] and later authors [3,46,48] defined and used the another descriptor, \( \text{ee}\%_{1/2} \). This is basically the \( \text{ee}\% \) value where the distribution function takes the value of 0.5. Enantiomeric excesses higher than this threshold value \((\text{ee}\%_{1/2})\) are expected to form with 50% probability. The following simple rule was deduced based on considerations using the properties of the normal distribution [3,49,52,55]:

\[
\text{ee}\%_{1/2} = \frac{0.6743}{\sqrt{N}}
\]

(22)

For better characterization, it would be good to calculate the expectation and standard deviation for \( \text{ee} \), which does not seem to have been done in any of the earlier works. Unpublished conjectures of the present author are the following:

\[
\langle \text{ee} \rangle_{e=0} = \frac{1}{N} \prod_{i=1}^{l} \frac{2i+1}{2i} \quad \langle \text{ee} \rangle \geq \langle \text{ee} \rangle_{e=0}
\]

(23)

\[
\sigma_{ee} = \sqrt{\frac{1}{N} + \frac{4}{N} - \langle \text{ee} \rangle^2}
\]

(24)

Turning now to stochastic non-CDS approaches to kinetics, Todorović et al. [62] investigated a slightly modified version of the classic Frank model [63]. Rather than considering the effect of individual molecules, the rate constants in the stochastic model were assumed to show stochastic fluctuations. Symmetric pairs of reactions producing different enantiomers of the same molecule were assumed to have uncorrelated fluctuations, which was the eventual origin of symmetry breaking within the deterministic rate equations. The differential equations were solved numerically with a fourth-order Runge–Kutta procedure using a pre-set finite step size with the rate constants varied according to a uniformly distributed random variables after each step. Conclusions were drawn based on regularities observed after several series of calculations started from identical, strictly non-chiral initial conditions. The results showed that the separation time, which is the shortest time after which the sign of the enantiomeric excess remains the same, was independent of the magnitude of fluctuations in the rate constant. In contrast, the parting time, defined as the moment at which the enantiomeric excess attains 1% of its limit for the first time, was a linear, decreasing function of the logarithm of the magnitude of the fluctuations in the rate constants. The drawback of this approach is that it assumes uncorrelated fluctuations in the values of rate constants which should be enantiomeric mirror-image pairs and the
origin of these fluctuations, which remained unspecified in the original work [62], may only be external. Consequently, this is a description of irreproducibility rather than inherent stochastic nature.

**Figure 3.** Probability of obtaining the favored enantiomer in excess in the case of asymmetric binomial distribution [54].

Cruz *et al.* also analyzed the effect of noise on the Frank model, but the random variables were introduced directly into the deterministic rate equations [64]. In fact, this is not very fortunate as it makes the questionable assumption that the rate equations themselves, and not the variables and parameters appearing in them, are subject to noise. Uncorrelated Gaussian white noise was included in the calculation of rates and the equations were solved by a fourth-order Runge-Kutta method with predetermined step size. With small noise amplitudes, no formation of significant enantiomeric excess was observed. When the noise amplitude was higher than a critical value, the enantiomeric excess converged to 100% in the numerical simulations. The article makes no clear statement in this regard, but implies that the added noises considered in the different rates were not scaled, *etc.* very different rates were subjected to noise of the same magnitude.

Rivera Islas *et al.* [65] studied three schemes related to the Frank model with deterministic simulations in order to study mirror-symmetry breaking. Surprisingly, they found spontaneous generation of sizable but unpredictable enantiomeric excesses starting from a completely nonchiral starting state. This effect primarily depended on the details of the numerical integrator used and was a consequence of the dynamic instability of the studied kinetic scheme. Micskei *et al.* made very similar observations [66] while studying an oscillatory model of the Soai reaction in a semibatch reactor. They proposed that the inherent numerical error of the integrator method might be a model of the statistical fluctuations. However, the present author feels that it is very questionable to draw chemical conclusions from effects that are admittedly caused by the insufficiency of the numerical method used to solve mathematical equations.

Hochberg and Zorzano [67] developed Langevin-type equations based on field theoretic techniques to describe inherent spatio-temporal fluctuations in the Frank model. Numerical simulations in two spatial dimensions showed that an enantiopure state can in fact form from strictly racemic initial conditions. The eventual origin of this symmetry breaking is the internal fluctuations in particle
density. Although this is a valid stochastic method, it still works based on continuous concentrations rather than discrete particle numbers. In a very well stirred reactor, the authors expected the round-off errors of the numerical calculations to be a decisive factor similarly the earlier results [65]. A later study from Hochberg and Zorzano also used the previous approach to study critical properties of the Frank Model [68]. An unstable, non-chiral fixed point, a pair of stable homochiral fixed points and racemic saddle point were characterized.

In more recent developments, Hochberg adopted an analytic perturbation method, which provides potentials associated with a wide class of stochastic partial differential equations, to introduce an effective potential for the Frank model [69,70]. This was also used to study the effects of external noise. Weak noise seemed to erase homochirality and stronger noise was needed to do the same when the system contained some chiral bias. Calculations were done in well mixed systems and also taking the effect of two dimensional diffusion into account.

Gleiser and Walker used limited stochastic elements in their study of a generalized autocatalytic model for chiral polymerization [71]. The system developed in this work was at some point coupled to a source of noise through a stochastic spatiotemporal Langevin equation to achieve small initial spatial fluctuations in the concentrations. However, the role of the stochastic approach and conclusions based on it are of secondary importance in this work.

5. The Continuous Time Discrete State Approach

5.1. General Characterization

The continuous time discrete state (CDS) stochastic kinetic approach is probably the closest to the currently accepted particle-based view of matter. The mathematics of CDS is authoritatively summarized in the book of Érdi and Tóth [44], and only the basic information necessary to understand the forthcoming discussion is given here. In CDS, the system is characterized by giving the numbers of each particle present. Instead of giving the concentrations as a function of time, the changes are described by giving the probabilities of all possible states as a function of time. A differential equation for the probability of every state is written based on the kinetic scheme. These Kolmogorov-like equations are referred to as the master equation of the process. The differential equations obtained in CDS, unlike in the deterministic approach, are all linear and can in principle be solved analytically. However, this analytical way of solving them is usually hindered by the fact that there are orders of magnitude more possible states than concentrations in the deterministic approach. Laplace transforms are often useful in the analytical solution of the master equation. The master equation can also be stated with matrix formalism. For this, an enumerating function, \( f \), is needed first, which gives a single and unique whole number to every possible state without repetitions or omissions. Therefore, if the number of possible states is \( M, f(c_1, c_2, c_3, \ldots, c_k) \) should take the values 1, 2, ..., \( M \) with \( c_1, c_2, c_3, \ldots, c_k \) being the molecule numbers for molecule types 1, 2, 3, ..., \( k \). If \( P \) is defined as a vector containing the probabilities of all states in the order of the values of the enumerating function, the master equation can be formulated as follows:

\[
\frac{dP}{dt} = \mathcal{Q}P
\]
Matrix $Q$ is the transition probability matrix (sometimes also called evolution matrix), the elements of which fall into two categories. Elements in the main diagonal are all negative and give the probability of the disappearance of each state. Elements outside the main diagonal in the $i$th row and $j$th column give the probabilities of forming state $i$ from state $j$. If all possible states are enumerated, all column sums in $Q$ are 0, therefore $Q$ is a singular matrix. For systems that have no elements of reversibility, etc., the same state cannot occur twice during the course of the process, it is often informative to calculate the probability ($Q$) of a given state ever occurring during the entire course of the reaction. It can be shown that $Q$ is related to function $P$ is follows:

$$Q(i) = \lim_{t \to \infty} P(i, t) + \int_{0}^{\infty} \sum_{j=1}^{M} Q(i, j) P(i, t) dt$$ \hspace{1cm} (26)$$

It should be noted that the CDS approach is superior to the deterministic kinetic approach because it allows for the particulate nature of matter without making any extra assumptions not already present in the deterministic approach. Every conclusion drawn with the deterministic approach can be reached as a limiting case of the CDS approach for very high particle numbers in very high volumes [48,72].

There is a simple relationship connecting stochastic ($\kappa$) and deterministic ($k$) rate constants, which is useful in comparing the two approaches. The formula only contains the Avogadro constant ($N_A$), the overall volume ($V$) of the system and the order of the rate constant ($i$):

$$\kappa = \frac{k}{(N_A V)^{i-1}}$$ \hspace{1cm} (27)$$

This relationship is actually the very origin of test 4 introduced in Section 4 to identify inherent stochastic nature in absolute asymmetric synthesis.

Despite its wide possible application range, the CDS kinetic approach is quite rare in the chemical literature. In addition to the issues connected to the way chemists are trained, a more objective reason for this lack of general interest is that CDS is much more computationally intensive than the deterministic kinetic approach. Although Equation 25 is a simple first-order differential equation, which can be solved analytically for any circumstances, transforming the solution into a form that is useful for comparing with experimental results is a quite separate and much more difficult issue. In addition, the concept of reproducibility is central in chemistry, so any hint of substantial random variations in experimental results is a solid reason for abandoning a project in most chemistry laboratories. One might actually wonder how much good stochastic experimental data could have remained unpublished through the last decades.

One minor exception from underreporting stochastic experimental observations is the field of enzyme catalysis. A stochastic solution to the Michaelis-Menten scheme has been described from different aspects [73–75] and experimental observations related to single-enzyme catalysis have also been published [76,77].

5.2. Simple Autocatalytic Schemes

Figure 4 shows a simple chiral autocatalytic scheme where molecule A is transformed to a chiral product B, usually using some sort of reagent(s) and/or catalyst(s). This scheme has been the subject of
extensive theoretical work in the last six years [35,37,38,78–80]. The enantiomers of B are $B_R$ and $B_S$. Homogeneous chiral autocatalysis is represented by two parallel reactions and rate expressions shown in Figure 4: a non-catalytic direct transformation with $v_{dir}$ rate and an autocatalytic step with $v_{autocat}$ rate. Similarly to the discussion of the binomial distribution in Section 4, a small asymmetry characterized by parameter $\varepsilon$ is considered in the direct transformation for generality.

In the mathematical discussion, $a$ will denote the number of A molecules, $r$ the number of $B_R$ molecules, and $s$ the number of $B_S$ molecules present. In the rate expressions, function $g(a)$ represents the rate of direct (not catalytic) pathway and it is only dependent on $a$. The scheme is also valid if the rate actually depends on additional reagents and catalysts because functions $g$ and $h$ only give the rate in a single kinetic run and may be different for different initial concentration(s) of reagent(s) and/or catalyst(s).

**Figure 4.** General symmetry breaking model with enantioselective autocatalysis.

At $t = 0$, the number of A molecules present is $N$, the number of $B_R$ and $B_S$ molecules present is 0, etc. there is no chiral material. Conservation of mass ensures that giving only the number of $B_R$ and $B_S$ molecules is sufficient to identify any possible state of the system unambiguously: $(r,s)$ will denote a state where the number $B_R$ molecules is exactly $r$, the number of $B_S$ molecules is exactly $s$, and consequently the number of A molecules is exactly $a = N - r - s$. The overall number of possible states ($M$) can be calculated easily:

$$M = \frac{(N + 1)(N + 2)}{2} \quad (28)$$

The enumerating function, if necessary, can be given as:

$$f(r,s) = \frac{(r + s)(r + s + 1)}{2} + r + 1 \quad (29)$$

The master equation of the process is the following:

$$\frac{dP(r,s,t)}{dt} = -P(r,s,t)[g(a) + h(a,r) + h(a,s)] + P(r-1,s,t)[0.5 + \varepsilon]g(a+1) + h(a+1,r-1)$$

$$+ P(r,s-1,t)[0.5 + \varepsilon]g(a+1) + h(a+1,s-1) \quad (30)$$
The total number of B molecules, \( b = r + s = N - a \), may also be needed for convenience in formulating certain equations. The initial state \((0,0)\) is certain at \( t = 0 \), therefore \( P(0,0,0) = 1 \), and \( P(r,s,0) = 0 \) is valid for every other state. As explained earlier, function \( Q(r,s) \) denotes the probability that the system goes through state \((r,s)\) at any time during the process. \( Q(0,0) = 1 \) holds because \((0,0)\) the certain initial state. The definition equations of the expectation for \( x_R \) in the final state and its standard deviation are as follows [38]:

\[
\langle x_R \rangle = \sum_{i=0}^{b} b_i Q(i,b-i) \tag{31}
\]

\[
\sigma_{x_R} = \sqrt{\sum_{i=0}^{b} \frac{i^2}{b^2} Q(i,b-i) - \langle x_R \rangle^2} \tag{32}
\]

Values of \( Q(r,s) \) can be given in a recursive formula, which shows that the final distribution only depends on the ratio of functions \( h \) and \( g \), which will be denoted \( \zeta(a,r) = h(a,r)/g(a) \).

\[
Q(r,s) = \left[ \frac{(0.5 + \varepsilon)g(a+1) + h(a+1,r-1)}{g(a+1) + h(a+1,r-1) + h(a+1,s-1)} \right] Q(r-1,s) + \right.
\]

\[
+ \left. \frac{[0.5 - \varepsilon)g(a+1) + h(a+1,s-1)]}{g(a+1) + h(a+1,r-1) + h(a+1,s-1)} Q(r,s-1) \right] \tag{33}
\]

First-order chiral autocatalysis used in a rather general sense means that the ratio of the catalytic and non-catalytic rates is directly proportional to the number of product molecules, etc. \( \zeta(a,r) = h(a,r)/g(a) = \alpha r \) irrespective of the value of \( a \). In this case, mathematical induction based on recursive Formula 33 can be used to obtain the values of \( Q \) [38]:

\[
Q(r,s) = \left( \frac{r+s}{r} \right)^{r-1} \prod_{j=0}^{r-s-1} (0.5 + \varepsilon + \alpha j) \prod_{j=0}^{s-1} (0.5 - \varepsilon + \alpha j) \prod_{j=0}^{r+s-1} (1 + \alpha j) \tag{34}
\]

The expectation for \( x_R \) and its standard deviation can also be given. The first is a constant and the second only depends the value of \( b (= r+s) \) in addition to the parameters [38]:

\[
\langle x_R \rangle = 0.5 + \varepsilon \tag{35}
\]

\[
\sigma_{x_R} = \sqrt{(0.5 + \varepsilon)(0.5 - \varepsilon) \frac{b^{-1} + \alpha}{1 + \alpha}} \tag{36}
\]

Finally, it can be shown that the discrete distribution given in Equation 34 converges to a continuous beta distribution with parameters \((0.5 + \varepsilon)\alpha^{-1}\) and \((0.5 - \varepsilon)\alpha^{-1}\) as \( N \) approaches infinity:

\[
P(x_R) = \frac{\Gamma(1/\alpha)}{\Gamma((0.5 + \varepsilon)/\alpha) \Gamma((0.5 - \varepsilon)/\alpha)} x_R^{(0.5+\varepsilon)/\alpha-1} (1-x_R)^{(0.5-\varepsilon)/\alpha-1} \tag{37}
\]

\[
D(x_R) = \frac{B(x_R,(0.5+\varepsilon)/\alpha,(0.5-\varepsilon)/\alpha)}{B((0.5+\varepsilon)/\alpha,(0.5-\varepsilon)/\alpha)} \tag{38}
\]
For \( \varepsilon = 0 \) (no preferential isomer), Equation 37 and Equation 38 reduce to a symmetric beta distribution. Asymmetric cases of this distribution can be analyzed further by introducing new variable \( R \) to denote the probability of forming \( B_R \) in excess over \( B_S \). The definition and subsequent mathematical consideration give the following formula for the general case [38]:

\[
R = \sum_{i=0}^{l} (Q(N-i,i) - Q(i,N-i)) = 2\varepsilon \sum_{i=0}^{l} \frac{Q(i,i)}{1 + 2\zeta(N-2i,i)}
\]

(39)

In Equation 39, \( l \) has the meaning that was defined in earlier Equation 20. The expectation for the enantiomeric excess in the final state can also be given:

\[
\langle ee \rangle = \sum_{i=0}^{l} \frac{N-2i}{N} (Q(N-i,i) + Q(i,N-i)) = 1 - \sum_{i=0}^{l} \frac{2i}{N} (Q(N-i,i) + Q(i,N-i))
\]

(40)

The description of the symmetric (\( \varepsilon = 0 \)) first-order autocatalytic system has been extended to initial conditions where some \( B_R \) and/or \( B_S \) molecules are present in the initial mixture [78] and have also been reproduced as a special case of a wider model [79]. The solution of the symmetric first-order autocatalytic system for initial conditions with non-zero \( B_R \) and/or \( B_S \) molecules gave the following \( Q \) function [78]:

\[
Q(r,s) = \left( \frac{N-r_0 + s_0}{r-r_0} \right)^{r-1} \prod_{j=r_0}^{r+s-1} (0.5 + \alpha j) \prod_{j=s_0}^{s-1} (0.5 + \alpha j)
\]

(41)

The enantiomeric excess and its standard deviation have been reported as [78]:

\[
\langle ee \rangle = \frac{l(N+1/\alpha)(r_0-s_0)}{N(r_0+s_0+1/\alpha)}
\]

(42)

\[
\sigma_{ee} = \frac{2}{N(r_0+s_0+1/\alpha)} \sqrt{\frac{(N+1/\alpha)(N-r_0-s_0)(r_0+0.5/\alpha)(s_0+0.5/\alpha)}{r_0+s_0+1/\alpha+1}}
\]

(43)

Turning now to higher-order autocatalysis, which can be represented generally by \( \zeta(a,r) = h(a,r)/g(a) = \beta f^{\zeta} \), the recursive definition is not easily transformed to an explicit form, but the probability of getting one enantiomer only, \( Q(r,0) \), can be given, which confirms that one of the enantiomers will be in overwhelming excess over the other for \( \xi > 1 \) and sufficiently large values of \( N \) [35,38]. However, higher-order autocatalysis in real examples does not necessarily lead to the formation enantiomeric excesses close to 100% because the convergence to the unique final distribution, unlike in the case of first-order autocatalysis, may not be very fast. For practically meaningful values of \( \beta \) and \( N \), the distribution does not approach the final limit and numerical calculations are necessary, which can be based on a combined stochastic deterministic approach. First, the appropriate values of \( Q(r,s) \) are calculated with the CDS approach for a relatively small value of \( b_0 \), e.g. 1,000 or 10,000 using the recursive formula given in Equation 33 in order to obtain the discrete distribution function [35]:

\[
D\left( \frac{r}{b_0} \right) = \sum_{i=0}^{r} Q(i,b_0) - i
\]

(44)
From the initial molar fraction of the R enantiomer \( (x_{R,i} = i/b_0) \), the expectation for the final molar fraction is calculated using the following deterministic differential Equation 35:

\[
\frac{dx_R}{db} = \frac{0.5 + \zeta(n-b,r)}{[1 + \zeta(n-b,r) + \zeta(n-b,b-r)(b+1)]} - \frac{x_R}{b+1}
\] (45)

For every initial value of \( x_{R,i} \), numerical integration of Equation 12 between the limits \( b_0 \) and \( N \) gives a final value for the molar fraction, \( x_{xR,i} \). The final distribution is given by \( F(x_{xR,i}) \). Figure 5 shows the experimental distribution of the S1 data set compared to theoretical predictions. The best fitting line (solid, red) is the one calculated assuming second-order autocatalysis [35]. The remaining two distributions shown are the one based on the best fit to first-order autocatalysis (symmetric beta distribution, blue dotted line) [37] and the combined beta distribution proposed without a model (green dashed line) in a recent work [36]. It was also pointed out that the differences in the fits are not statistically significant [35], but the primary reason for this might be the relatively low number of measured points.

**Figure 5.** Fit of different theoretical distributions (\( D \): distribution function) to data set S1. Solid red line: second-order autocatalysis [35], blue dotted line: first-order autocatalysis [37], green dashed line: combined beta distribution [36].

Finding explicit formulas for function \( R \) and the expectation for the enantiomeric excess in the case of higher-order autocatalysis seems to be very difficult. However, the predictions of the CDS approach can still be meaningfully assessed by proving certain limiting formulas for different parameters. A higher limit can be obtained for the value of \( R \) and consequently for enantiomeric excess [38]:

\[
R < 2^\epsilon + \frac{2\epsilon}{\sqrt{\pi}} \int_0^{0.5} \frac{d\xi}{\sqrt{\xi + 2} \xi^{0.5}} + 2^\epsilon \beta e^{(0.5 - \epsilon)} \] (46)

Furthermore, an upper limit can be given for the ratio of \( Q(b,0) \) and \( Q(0,b) \) [38]:

\[
\frac{Q(b,0)}{Q(0,b)} < e^{4\epsilon/(0.5 - \epsilon) + 2(\xi - 1)e/\beta}
\] (47)
Figure 6 displays some theoretical fits to experimental distributions S2, which has been confirmed to show significant asymmetry [38–39]. The best fitting curve here is the asymmetric second-order autocatalysis (red solid line) [38], whereas the combined beta distribution (green dashed line) also gives acceptable fit but with many more parameters [36]. The final curve (blue dotted line) is the best fitting symmetric second-order autocatalytic distribution, it is shown for comparison purposes as statistic analysis clearly rejects any symmetric fit [39].

**Figure 6.** Fit of different theoretical distributions (\(D\): distribution function) to data set S2. Solid red line: second-order autocatalysis with a favored isomer [38], blue dotted line: second-order autocatalysis without a favored isomer [35], green dashed line: combined beta distribution [36].

The results of the CDS approach with respect to the possibility of inherent asymmetry are best demonstrated using some arbitrarily defined conditions, which identify certain regions in the parameter space of the model [38]. When the standard deviation of the expectation of \(x_R\) is larger than 0.01, conditions are said to fall into the stochastic region [38] of the model, where the final distribution of enantiomers shows average fluctuations larger than 1%. Consequently, the deterministic approach is unsuitable for describing the system.

The term *de lege* region [38] is only used for \(\varepsilon > 0\) and conditions under which the probability of obtaining the \(B_R\) enantiomer in excess (R) is higher than 0.9. In this region, the inherent asymmetry determines which enantiomer accumulates during amplification. The ‘significant \(B_R\) excess region’ [38] is again defined for \(\varepsilon > 0\) to denote cases where the difference between the expectation of \(x_R\) and 0.5 is larger than the standard deviation of \(x_R\). This is very close to the *de lege* region.

Finally, the ‘high \(ee\) region’ [38] was defined as the part of the parameter space where the expectation of the enantiomeric excess formed is higher than 0.9. This is characteristically different from the *de lege* region, as it does not matter which of the two enantiomers is in excess.

With the estimated \(\varepsilon = 10^{-17}\) [54,60,61], the beta distribution is not significantly different from the one obtained with \(\varepsilon = 0\). Therefore, the parity violating energy difference is without significant consequence in the case of first-order autocatalysis. The different regions characteristic of second-order autocatalysis (\(\xi = 2\)) are shown in Figure 7. The *de lege* region and the high \(ee\) region have a
common area but only when the initial number of A molecules is above $10^{50}$. This is an unreasonably large number: $10^{50}$ molecules of alanine have approximately the same mass as planet Earth. It is also noticeable that the stochastic region does not contain the high ee region fully. No significant stochastic effects arise if $\beta < 2 \times 10^{-67}$, whereas a suitable $N$ can be found within the high ee region for every value of $\beta$. It was estimated that the de lege and high ee regions cannot have common area if the initial amount of substance for reactants is smaller than 100 Gmol, no matter what the values of $\xi$ and $\beta$ are. Therefore, if present estimates [60] of the parity-violation energy difference are true, this phenomenon could have no consequence for the origin of biological chirality. Data set S2 measured in the Soai reaction gives an acceptable-looking fit to second-order autocatalysis and $\varepsilon = 6.6 \times 10^{-11}$ [38], which is about seven orders of magnitude higher than could be expected based on the theoretically predicted parity violation energy difference between the enantiomers of the product of Equation 1 [61].

Some efforts have also been made to use the CDS approach to describe the stochastic phenomena in the reaction times as well. No analytical solutions were found here, but numerical studies [35] showed that even first-order autocatalysis can interpret the random fluctuations observed experimentally in the reaction between chlorite ion and iodide ion [21].

**Figure 7.** Different regions in the parameter space for second-order autocatalysis as a function of parameters $N$ and $\beta$ [38]. (1: high ee region, 2: stochastic region, 3: de lege region).

5.3. Schemes with Recycling

Saito, Sugimori, and Hyuga considered the full stochastic kinetic model based on Scheme 1 but the role of reverse reactions, called recycling, was also considered [79,80]. The deterministic equivalent of recycling was later shown to violate the second law of thermodynamics and the principle of detailed equilibrium [81–83]. Still, these results deserve closer attention as recent work pointed out that the mathematical equations used in recycling are identical to an open-system scenario with inflow and outflow [84].

The earlier work of Saito et al. [79] studied a scheme that can be thought of as a symmetric version of the one given in Figure 4 with $h(a,r) = \kappa_1 a r + \kappa_2 a r^2$ but instead of $g(a)$ implying an irreversible process, a reversible analog is used with $g(a,r,s) = \kappa_0 a - \lambda r - \lambda s$. The master equation is:
\[
\frac{dP(r,s,t)}{dt} = -P(r,s,t)(2\kappa_0 a + \kappa_1 ar + \kappa_1 as + \kappa_2 ar^2 + \kappa_2 as^2 + \lambda r + \lambda s) + \\
+ P(r-1,s,t)\left[\kappa_0 (a + 1) + \kappa_1 (a + 1)(r - 1) + \kappa_2 (a + 1)(r - 1)^2\right] + \\
+ P(r,s-1,t)\left[\kappa_0 (a + 1) + \kappa_1 (a + 1)(s - 1) + \kappa_2 (a + 1)(s - 1)^2\right] + \\
+ P(r+1,s,t)\lambda (r + 1) + P(r,s-1,t)\lambda (s - 1)
\]

In this system, none of the states is final. Therefore, \( Q \) functions carry no meaning and the \( S(r,s) \) stationary probabilities should be defined as the final, time-independent probability values for each state:

\[
S(r,s) = \lim_{t \to \infty} P(r,s,t)
\]  

The following formula was derived starting from master Equation 48 [79]:

\[
S(r,s) = \Xi \frac{N!}{r!s!(N - r - s)!} \lambda^{N-r-s} \prod_{j=0}^{r-1} (\kappa_0 + \kappa_1 j + \kappa_2 j^2) \prod_{j=0}^{s-1} (\kappa_0 + \kappa_1 j + \kappa_2 j^2)
\]

In Equation 50, \( \Xi \) is a normalizing constant and the appropriate products shown should simply be replaced by a number 1 if \( r = 0 \) or \( s = 0 \). Continuous approximations of the final distributions have also been given. Numerical integration of the master equation was carried out for \( N = 100 \) to study the evolution of the probability distribution. Figure 8 shows the evolution of the statistical distributions for the case of second order autocatalysis [79]. An eigenvalue analysis of the evolution matrix \( \Omega \) was also presented [79].

**Figure 8.** Contour plot of the time evolution of statistical distributions in a scheme with second-order autocatalysis and recycling [79]. Parameters: \( \kappa_2 = 0.001 \times \kappa_0, \lambda = \kappa_0, N = 100, t \kappa_0 = 0.0536 \) (green), 0.3362 (red), 0.7915 (blue), \( \infty \) (black).
Finally, this work [79] introduced a random walk approach to CDS kinetics on a square lattice \((r,s)\) in a triangular region \(0 \leq r, s, (r + s) \leq N\) for studying the system without recycling. The random walker can only make directed walks to the \(R\) and \(S\) directions: the possible steps are \((r,s) \rightarrow (r+1,s)\) and \((r,s) \rightarrow (r,s+1)\). Based on the transition probabilities in the master equation, the walker on site \((r,s)\) stays on the site for a waiting time interval of [79]:

\[
\tau(r,s) = \frac{1}{[2\kappa_0 + \kappa_1(r + s) + \kappa_2(r^2 + s^2)](N - r - s)}
\]

(51)

After the waiting time of \(\tau\), the jump is in either \(R\) or \(S\) direction with the respective probabilities of \(p_R\) and \(p_S\) [79]:

\[
p_R(r,s) = \frac{\kappa_0 + \kappa_1 r + \kappa_2 r^2}{2\kappa_0 + \kappa_1(r + s) + \kappa_2(r^2 + s^2)}
\]

(52)

\[
p_S(r,s) = \frac{\kappa_0 + \kappa_1 s + \kappa_2 s^2}{2\kappa_0 + \kappa_1(r + s) + \kappa_2(r^2 + s^2)}
\]

(53)

The essence of the random walk model and Equation 52 and Equation 53 quite analogous to the use of \(Q\) functions in some respect: the time dependence is separated from the variables of \(P\). The recursive formula for \(Q\) values given in (33) can be thought of as a sort of random walk. The significance of this observation is that a full-scale Monte-Carlo method to the CDS approach could be based on a number of random walk procedures and used for approximating distributions of enantiomeric excesses or reaction times even if the chemical model is quite complicated and direct analytical or numerical solutions of the master equation is hindered by the very large number of states possible.

The same three authors later published a very similar analysis on a scheme which does contain any autocatalysis [80]. They showed that symmetry breaking may occur based on recycling alone. The considered reactions and rate equations are as follows:

\[
A \rightarrow B_R \ \upsilon = \kappa_0 \delta \alpha
\]

(54)

\[
A \rightarrow B_S \ \upsilon = \kappa_0 \delta \beta
\]

(55)

\[
R + S \rightarrow 2A \ \upsilon = \kappa_0 \delta \gamma
\]

(56)

However, the model used contradicts the principle of microscopic reversibility [67]. A flow-through system analog [84] does not seem particularly reasonable in this case as the feed should contain a perfectly racemic mixture of \(B_R\) and \(B_S\) so that the ‘flow reaction’ would replace Process 54 and Process 55. Therefore, this could only be a thermodynamically acceptable model of a flow-through reactor with pre-existing chiral (but racemic) material and the reaction itself would destroy some of the chiral material to create significant enantiomeric excesses, which may not be strongly relevant for the origin of biological chirality. But again, the example showed that use of the CDS approach is necessary to understand phenomena which seem impossible within the realm of the deterministic kinetic approach.
5.4. The Frank Model

Frank proposed his classic model in 1953 to give a theoretical mechanism that could lead to asymmetric synthesis [63]. This model has been the subject of more studies than can be cited here, but only a few of them had explicit stochastic elements [62,64,67−80,84]. The non-CDS studies [62,64,67−70] have already been discussed. The latest work [85] used the CDS approach primarily with numerical calculations for relatively low molecule numbers (100−1000) and without major success in finding explicit formulae for chemically more meaningful amounts of substances [85].

Figure 9. The Frank model.

The model is shown in Figure 9: it basically comprises a first order version of the simple autocatalytic model shown in Figure 4 with the addition of an extra step, termed mutual antagonism by Frank [63], whereby the two enantiomers $B_R$ and $B_S$ react with each other to give nonchiral product C. Possible states can be identified by giving the numbers of molecules $A$ ($a$), $B_R$ ($r$) and $B_S$ ($s$) as conservation of mass ensures $c = N - a - r - s$ for the number of C molecules. In this system, the overall number of possible states ($M$) is a third-order polynomial of $N$:

$$M = \binom{N+3}{3} = \frac{(N+3)(N+2)(N+1)}{6}$$ (57)

A convenient enumerating function is:

$$f(a,r,s) = \frac{a^3}{6} + \frac{aN(a+N)}{2} - a^2 + 2aN + \frac{11a}{6} + \frac{r^2}{2} + r(n-a) + \frac{3r}{2} + s + 1$$ (58)

The master equation can be written as:

$$\frac{dP(a,r,s,t)}{dt} = -\{2a\kappa_u + ar\kappa_c + as\kappa_c + rs\kappa_d + (n-a)\kappa_f\}P(a,r,s,t) +$$

$$+ \{(a+1)\kappa_u + (a+1)(r-1)\kappa_c\}P(a+1,r-1,s,t) +$$

$$+ \{(a+1)\kappa_u + (a+1)(s-1)\kappa_c\}P(a+1,r,s-1,t) +$$

$$+ \{(r+1)(s+1)\kappa_d\}P(a,r+1,s+1,t) + (r+1)\kappa_f P(a-1,r+1,s,t) +$$

$$+ (s+1)\kappa_f P(a-1,r,s+1,t) + (n-a-r-s+1)\kappa_f P(a-1,r,s,t)$$ (59)
Parameter \( \kappa \) here corresponds to the flow rate, the equation itself also describes a closed system without flow if \( \kappa = 0 \) is set. The expectation for enantiomeric excess in the reactor at any time is given as usual [85]:

\[
\langle ee \rangle(t) = \sum_{\text{all } M \text{ states}} \left| \frac{r-s}{r+s} \right| P(a,r,s,t)
\] (60)

However this traditional enantiomeric excess can be grossly misleading, as also pointed out in a deterministic analysis of closed systems [83]. Therefore, another descriptor, the yield-adjusted enantiomeric excess (\( E \)) was introduced, which is formally the product of \( ee \) and the overall yield of the chiral product [85]:

\[
\langle E \rangle(t) = \sum_{\text{all } M \text{ states}} \left| \frac{r-s}{N} \right| P(a,r,s,t)
\] (61)

The stationary (or, in closed systems, final) values of these descriptors are defined as [85]:

\[
\langle ee \rangle = \lim_{t \to \infty} ee(t)
\] (62)

\[
\langle E \rangle = \lim_{t \to \infty} E(t)
\] (63)

The closed system proceeds to a limited number of final states, from which no further change can arise. Obviously, \( Q(N,0,0) = 1 \) holds because \( (N,0,0) \) the certain initial state. There can be no A molecules in final states, and \( B_R \) and \( B_S \) molecules cannot coexist, either. The following equations have been proven [85]:

\[
\langle ee \rangle = 1 - Q(0,0,0) \approx 1
\] (64)

\[
\langle E \rangle = \frac{1}{N} \sum_{i=1}^{N} 2i Q(0,i,0)
\] (65)

The usual recursive formula can be derived for \( Q(a,r,s) \) values [85]:

\[
Q(a,r,s) = \frac{((a+1)\kappa_u + (a+1)(r-1)\kappa_c)Q(a+1,r-1,s,t)}{2(a+1)\kappa_u + (a+1)(r+s-1)\kappa_c + (r-1)s\kappa_d} +
\]

\[
\frac{((a+1)\kappa_u + (a+1)(s-1)\kappa_c)Q(a+1,r,s-1,t)}{2(a+1)\kappa_u + (a+1)(r+s-1)\kappa_c + (s-1)\kappa_d} + \frac{(r+1)(s+1)\kappa_d Q(a,r+1,s+1)}{2a\kappa_u + a(r+s+2)\kappa_c + (r+1)(s+1)\kappa_d}
\] (66)

With an appropriate calculation sequence, all \( Q(a,r,s) \) values can be computed numerically conveniently for \( N \) values up to 1,000 [85]. The proved symmetries of the master equation and the solution can be used to deduce a symmetry-reduced version of the master equations, which improves the speed of computation. The master equation is homogeneous in the stochastic rate constants and, consequently, parameters can be scaled with one of the \( \kappa \) values with no loss of information. Figure 10 shows the probability distributions of final states calculated for \( N = 200 \) at several different \( \kappa_d/\kappa_u \) values with \( \kappa_d/\kappa_u = 0.01 \) fixed. The distribution at low \( \kappa_d/\kappa_u \) values (0.01 and lower) tends to the beta distribution given in Equation 37 and Equation 38 because the reaction responsible for mutual antagonism only occurs after all A molecules are converted to B. At higher \( \kappa_d/\kappa_u \), the maximum at
\( r = 0 \) splits into a symmetric pair, and this bimodal distribution is unchanged when \( \kappa_d/\kappa_u \) is increased above 10. At this extreme, \( B_R \) or \( B_S \) can only coexist for very short time periods [85].

**Figure 10.** Final probability distributions in the Frank model in a closed system at \( \kappa_d/\kappa_u = 0.01 \). Numbers displayed inside the graph show the value of \( \kappa_d/\kappa_u \) for each distribution shown [85].

Studying the effect of increasing the overall number of molecules revealed that there are two separate ways of increasing molecule numbers in stochastic kinetic models: at constant concentrations and increasing volume and at constant volume and increasing concentrations [85]. The first case showed the effects expected based on Kurtz’s theorem [48].

There is a substantial difference between the CDS descriptions of the Frank model in closed and flow-through systems [85]. Open-flow CDS approaches are very rare in the literature, but not without precedent [86]. The effect of flow can be thought of as a first-order process which replaces all molecules in the reactor with those in the feed [85]. In a flow-through process, there are no final states and each state can arise from or give rise to another state(s). Stationary conditions, in which none of the probabilities change any more, play a similar role in flow-through reactors as final states do in closed systems. The stationary probabilities, which are conceptually very different from the stationary concentrations used in the deterministic approach to flow-through reactors, can be calculated for all states by writing 0 on the left side of Equation 59 [85]:

\[
0 = -\{2a\kappa_u + ar\kappa_c + as\kappa_c + rsk_d + (n-a)\kappa_f \}S(a,r,s) + \\
+ \{(a+1)\kappa_u + (a+1)(r-1)\kappa_c \}S(a+1,r-1,s) + \\
+ \{(a+1)\kappa_u + (a+1)(s-1)\kappa_c \}S(a+1,r,s-1) + \\
+ \{(r+1)(s+1)\kappa_f \}S(a,r+1,s+1,t) + (r+1)\kappa_f S(a-1,r+1,s) + \\
+ (s+1)\kappa_f S(a-1,r,s+1) + (n-a-r-a+1)\kappa_f S(a-1,r,s)
\]

(67)

In this way, the original set of \( M \) simultaneous differential equations is transformed into \( M \) homogeneous linear algebraic equations, any one of which can be obtained by summing the remaining \( M-1 \) equations. One more independent equation can be written based on the fact that all the states have
been listed: the sum of all probabilities is one. These equation have a unique set of solutions, etc. the stationary probabilities are independent of initial conditions. Although the algebraic equations are linear, a special iteration method had to be developed to solve them for $N = 100$ ($M = 176851$) [85]. Further improvements in computational power are unlikely to help the numerical problems significantly, which are not primarily limited by computation time but available memory. Figure 11 shows the expectation for $E$ and $ee$ as a function of both $\kappa_d/\kappa_f$ and $\kappa_c/\kappa_f$ at fixed $\kappa_u/\kappa_f$ and $n = 100$. The front left portions of both graphs provide examples when considering $ee$ only would result in questionable conclusions. Relatively high values of $\kappa_c$ (etc. efficient enantioselective autocatalysis) are the key to getting high $E$ or $ee$ values. For low values of $\kappa_d$ (mutual antagonism), it has no influence on the course of the overall process.

An unexpected feature of Figure 11 is the ‘valley’ or ‘gorge’ at about $\kappa_d/\kappa_f = 10^{-4}$ and $\kappa_c/\kappa_f > 1$, where intermediate values $\kappa_d/\kappa_f$ can lead to an overall negative effect on the expectation of both $E$ and $ee$ because the autocatalyst is removed from the system, whereas higher $\kappa_d/\kappa_f$ values the coexistence of BR and BS is impossible, which again favors high expectations for $ee$ and $E$ [85]. It was also shown that a flow-through system is not bound by Kurtz’s theorem [48].

**Figure 11.** Expectation of the yield-adjusted and conventional enantiomeric excess for the Frank model in an open system with $\kappa_u/\kappa_f = 1$ and $N = 100$ [85].
6. Conclusions and Outlook

The resounding single conclusion of this review is that the use of stochastic models is inevitable in interpreting the origins of biological chirality. The CDS approach, used with some mathematical resourcefulness, can provide a very useful tool in interpreting absolute asymmetric reactions and the enantiomeric distributions observed in them. The CDS approach is well developed and is readily adopted for enantioselective autocatalysis, the simplest examples of which have now been studied with mathematical rigor.

The Soai reaction and other examples of absolute asymmetric synthesis have much more complex reaction mechanisms [87–89] than the ones studied by the CDS approach. Even with these unreasonably simple models, some success was achieved in predicting measured statistical distributions. More complex kinetic schemes, relying on the results of mechanistic studies, need to be studied in the future by the CDS approach for deeper understanding. A full-scale Monte-Carlo approach that extends the principles of the random-walk model already used in two papers [79–80] seems to have large computational potential for studying more complex reaction mechanism. In experimental work, the kinetics of the Soai reaction should be thoroughly studied under absolute asymmetric conditions and the expected random effects must be characterized.

Finally, it should also be noted that claims for thermodynamic equilibrium as a driving force of symmetry breaking exist in the relevant literature [90,91]. It is necessary to further substantiate these claims based on a stochastic approach to thermodynamics. Unlike in the case of CDS, the use of such an approach, under the name statistical thermodynamics, is often covered in detail in university physical chemistry courses and textbooks.

References and Notes


© 2010 by the author; licensee Molecular Diversity Preservation International, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).