CHAPTER 2: SSB AT MOLECULAR SCALES – PURE AUTOCATALYSIS

2.1 Introduction

Chemical reactions in solution occur when the chemically reacting molecules diffuse and collide with each other. In addition to this spontaneous formation of products, there can be enzymes which enhance the probability of the reaction by lowering the activation barrier. Many biochemical reactions are autocatalytic, with the product catalyzing its own production.

In this chapter, we discuss the interplay between chemical fluctuations and autocatalysis in driving a mixed (racemic) chemical systems towards spontaneous chiral symmetry breaking (CSB), thereby obtaining reaction products with specific chirality.
2.2 Models of autocatalysis

Consider a well stirred chemical reaction of the form described in Fig. 2.1, where the reactants A and B combine to form products R and \( R^* \) (chiral enantiomers or indeed any binary attribute). The forward reaction is purely autocatalytic with rates \( k_1R \) and \( k_1R^* \), respectively – existing enantiomers act as a “local field” favouring their own production. On the other hand, the reverse reactions can be non-autocatalytic with identical rates \( k_2 \). We stress that the whole reaction system is perfectly symmetric. Note that this model is much simpler than the Frank model described in Chapter 1, in that there is no explicit antagonism.

![Diagram of Models with linear autocatalysis](image)

Figure 2.1: Models with linear autocatalysis (\( k_1R, k_1R^* \)): (a) Achiral reactants A and B form the chiral enantiomers, R and \( R^* \) (b) Achiral reactant A produces chiral enantiomers R and \( R^* \)

A simple linear model of autocatalysis is \( k_1(R) = k_1P(N_R) \) and \( k_1(R^*) = k_1P(N_{R^*}) \), where \( P(N_R) (P(N_{R^*})) \) is the probability of encountering \( R(R^*) \) in the forward reaction.

2.3 Mean-field analysis

The mean-field autocatalytic reaction-diffusion equations of Fig. 2.1(a), valid when the local fluctuations in the concentration \( n \) of each species is much
smaller than the mean $\bar{n}$, $\sqrt{\bar{n}^2 - \bar{n}^2} \ll \bar{n}$, follows mass action kinetics and is given by,

\begin{align}
\frac{\partial n_A}{\partial t} &= -(k_1 (n_R + n_{R^*})n_A n_B + k_2 (n_R + n_{R^*}) + D_A \nabla^2 n_A \quad (2.1) \\
\frac{\partial n_R}{\partial t} &= k_1 n_A n_B n_R - k_2 n_R + D_R \nabla^2 n_R \quad (2.2) \\
\frac{\partial n_{R^*}}{\partial t} &= k_1 n_A n_B n_{R^*} - k_2 n_{R^*} + D_{R^*} \nabla^2 n_{R^*} \quad (2.3)
\end{align}

Note that the rates of reactions, $k_1$ and $k_2$ have units of $m^6 s^{-1}$ per mole$^2$, and $s^{-1}$ respectively.

These equations can be solved exactly in the well-stirred limit, to obtain $dn_R/dn_{R^*} = n_R/n_{R^*}$. This implies that with symmetric initial conditions, $n_R(0) = n_{R^*}(0)$, there can be no CSB, i.e., $n_R(t) = n_{R^*}(t)$ at all times.

A fixed point analysis of (2.1)–(2.3) shows that there is a line of neutral stability (to linear perturbations) with $n_A = n_B = \sqrt{k_2/k_1}$ — marginal fixed points corresponding to both chiral symmetric and asymmetric states. The relative enatiomeric concentration $\phi \equiv n_R - n_{R^*} = 0$. An arbitrary perturbation from the chiral symmetric fixed point generated by the slightest initial stereo-preference, will remain as is (i.e., unamplified) and will not give rise to a complete CSB. This is in contrast to (21) – the zero enantiomeric excess is an unstable equilibrium and an initial asymmetry, if any, is exponentially amplified (Fig. 2.2).

We next study the effect of including an additive or a multiplicative noise in the mean-field equations. Such sources of noise arise in an open chemical system through spatiotemporal fluctuations in the number or temperature or indeed any other ‘fast’ variable that couples to the chemical rates. The additive (multiplicative) noise is distributed independently and uniformly with zero (nonzero) mean and variance chosen such that the concentrations
Figure 2.2: Phase portrait of (left to right): (i) Frank’s model with an unstable fixed point at the origin and a saddle node at \((\frac{k_1}{k_2}, \frac{k_2}{k_1})\) (= 10,10 here). The pink line corresponds to trajectories of symmetric initial conditions. Any initial bias is exponentially amplified, converging to one of the asymptotes \(R=0\) or \(R^*=0\) (ii) Our autocatalytic model (a) with a line of neutrally stable fixed points. Note that any bias in initial condition stays unamplified.

are non-negative at all times. We have numerically checked that such an implementation of stochasticity does not lead to complete CSB starting from symmetric initial conditions. Indeed in multiplicative noise models, the zero relative enantiomeric concentration \(\phi \equiv n_R - n_{R^*} = 0\) is an absorbing state. As an aside, we note the difference between this and the role of multiplicative noise in (36).

An important feature of mean-field chemical kinetics is that in every interval \(\Delta t\), all reactions take place simultaneously. This feature is retained when we introduce stochasticity as above – thus, this kind of stochasticity is incapable of generating CSB from symmetric initial conditions.
2.4 Stochastic formulation of chemical reactions in a well-stirred system

It is clear that we can no longer deal with deterministic chemical dynamics when the fluctuations in the number of reactant molecules is of the same order of magnitude as the number of molecules. As noisy biochemical reactions are fundamentally discrete stochastic events, one needs to treat the reaction-diffusion kinetics as a Markov process described by chemical master equation (37).

We construct a local coarse-grained ‘volume’ ξₗ (where d = 2, 3) by noting that at scales smaller than ξ ∼ \sqrt{Dt_r}, where D is the typical diffusion coefficient and t_r is the typical reaction time, the chemical system may be considered well-stirred.

2.4.1 Chemical Master equation

Consider a well-mixed dilute system of chemically interacting molecules, such that the following approximations are valid (38):

1. The position and velocity of all the molecules are statistically independent; they are uniformly distributed in space and obeys Maxwell-Boltzmann distribution for velocity as it is in thermal equilibrium at a temperature T.

2. Probability of more than two-body collision is negligible and hence typically a reaction per unit time interval happens.

With this we can write the Master equation in the full distribution function as a chemical master equation in terms of a reduced distribution function, where particle momenta have been integrated out. The reaction
processes, corresponding to our model mechanism, are described by,

$$\frac{\partial P(N, t)}{\partial t} = \sum_{N'} W_{NN'} P(N', t) - W_{NN} P(N, t)$$  \hspace{3cm} (2.4)$$

where $P(N, t)$ is the probability that the system is in state $N \equiv N^{(i)} \equiv \{N_A^i, N_B^i, N_R^i, N_{R^*}^i\}$ ($N^i$ is the number of molecules of the $\alpha$-species in box $i$) at time $t$.

$W_{NN'}$ is the transition probability per unit time to go from state $N' \rightarrow N$.

The transition matrices $W$ consist of sums of terms, obeying detailed balance, corresponding to the chemical reactions with ‘mass action’-like kinetics. For instance, for the forward $R$-reaction,

$$W_{NN'} = k_1 N_A^i N_B^i N_R^{i*} \delta_{N_A,i} N_A^{i+1} \delta_{N_B,i} N_B^{i+1} \delta_{N_R,i} N_R^{i-1} \delta_{N_{R^*},i} N_{R^*}^{i*},$$  \hspace{3cm} (2.5)$$

where we have suppressed the index $i$ for clarity. The reaction propensity $k_1 N_A^i N_B^i N_R^{i*}$ is the probability per unit time of this reaction occurring.

In general, this master equation is hard to solve analytically; we therefore obtain $P(N^{(i)}), t)$ through a Monte Carlo simulation technique known as the Gillespie algorithm (39).

### 2.5 Stochastic simulation: the Gillespie (direct) algorithm

The Gillespie algorithm is an event-based update algorithm, which generates trajectories $\{N^{(i)}(t)\}$ in exact accord with the probability distribution $P(N^{(i)}), t)$. The stochasticity in molecular encounters are taken care of appropriately with Poisson statistics for the waiting time distribution. In Gillespie’s Direct algorithm, typically one asks the following question: What
is the probability that the next reaction is $\mu$ and that it happens at time $\tau$.

The reaction propensity is given by,

$$P(\mu, \tau) = a_\mu \exp(-\tau \sum_j a_j) \, d\tau$$

The basic assumption is that the probability of choosing a reaction is completely independent of the time of next reaction, hence integrating $P(\mu, \tau)$ over all $\tau$ from $0$ to $\infty$ one obtains,

$$P(\mu) = \frac{a_\mu}{\sum_j a_j}$$

The time step for next reaction is chosen, by summing $P(\mu, \tau)$ over all $\mu$, from the following exponential distribution,

$$P(\tau) \, d\tau = \left(\sum_j a_j\right) \exp\left(-\tau \sum_j a_j \, d\tau\right)$$

This overcomes the computational inaccuracies and instabilities due to finite time step approximation in numerical methods.

This Direct method takes a time proportional to the number of reactions. There are better methods, e.g., the Next reaction method (40), that take a time proportional to logarithm of the number of reactions. Here we will use the easily implementable Direct method.

### 2.6 Chiral symmetry breaking (CSB)

We choose a perfectly symmetric initial condition with \{\(N_A^i = N_B^i, N_R^i = N_R^{i*}\)\} in a single box. The parameter values are chosen as $c_1 = 10^{-3} \text{s}^{-1}$, $(k_1 = c_1 \xi^4)$ and $k_2 = 10^{-1} \text{s}^{-1}$. Our qualitative results remain unaltered.
when we vary the rates $c_1 = 10^{-3} - 10^{-5}$ and $k_2 = 10^{-1} - 10^{-2}$ (which merely changes the time and space scales). We measure time and space in units of $1/k_2$ and $\xi$, respectively. Averages denoted by $\langle \ldots \rangle$ are over 10 realisations of random number seeds, further the temporal behaviour of various observables are smoothened by binning over a time interval of $t = 5$.

We define a local order parameter, the relative enantiomeric excess,

$$\Phi = \left\langle \frac{N_R - N_{R^*}}{N_R + N_{R^*}} \right\rangle$$ (2.6)

In spite of perfectly symmetric initial data, the stochastic evolution of the populations of $R$ and $R^*$ in a typical box, shows complete CSB beyond a time scale $\tau$ (Fig. 2.3(a)) – leading to a complete (100 %) selection of one of the degenerate steady-states (either $R = 0$ or $R^* = 0$). The enantiomer (say $R$) takes over the entire population, $\langle N_R \rangle \propto 1 - \exp(-\beta t)$, with a rate that is an increasing function of $k_1$. The average number of $A$ (and $B$) molecules decrease with time to an asymptote which is a weak function of initial conditions. Hence we achieve symmetry breaking spontaneously, without invoking (i) bias in initial condition (ii) antagonistic reactions, unlike (21).

At first this result seems surprising, since, fluctuations are known to destroy long-range order in systems at equilibrium (41). However we do know that noise can also induce second-order transition through a mechanism that involves the interplay of multiplicative noise, nonlinearity and diffusion (42, 43). Further, in systems far from thermodynamic equilibrium (e.g. Lotka-Volterra system and glycolysis), autocatalysis can lead to the creation of temporal order resulting in stable limit cycle and spatial (e.g. Brusselator) order (44).

The CSB in our system arises from the interplay of autocatalysis and
Figure 2.3: (a) Number of reactants versus time for a given set of initial data: A and B saturate to a constant population, there is a definite enantiomeric selection between R and R*, which results in CSB. (b) The time evolution of the relative enantiomeric excess $\langle \Phi \rangle$, with error bars from different noise realisations.
number fluctuation inherent in chemical reactions, and leads to exponential amplification of chiral enantiomeric excess; $\Phi = \pm 1$ (45, 46). It is evident that once the system is completely locked into a degenerate state (say $\Phi = 1$), there is no way it can sample the other degenerate configuration ($\Phi = -1$) as $R^* = 0$.

![Scaling of the mean time for chirality discrimination](image)

Figure 2.4: Scaling of the mean time for chirality discrimination with total number of reactants for different initial data, showing $\langle \tau \rangle \propto N$. Inset shows the distribution $P(\tau) \sim \tau \exp(-\tau/\tau_0)$ for $(N_A(0) = 100, N_R(0) = 10)$, averaged over 500 realisations of the noise.

By extending our model to include nonlinear autocatalysis, as in (28), we find that this tendency for complete CSB only gets enhanced. Importantly, the CSB that we report is fluctuation induced, and does not arise in the mean field theory.

Note that in our model, there is no need to invoke specific antagonism, i.e., the presence of one species need not destroy the other, unlike the catalyst-antagonist chemical reactions (21). The survival of either species is decided purely by chance – every realisation has an equal probability of forming R or
Moreover, unlike all the other theoretical models and their experimental realisations (2), our model achieves complete amplification – the fractional enantiomeric excess is a 100%!

Thus the steady state value of the relative enantiomeric excess $\Phi$ is either $+1$ or $-1$ (Fig. 2.3(b)). One of these steady-states is dynamically chosen in every realisation, resulting in a bimodal distribution, ($\Phi = +1$ or $-1$, binary decisions). As in the dynamics towards any symmetry broken state (e.g. a quench from the paramagnet to the ferromagnet), an average over different random seeds averages over the two degenerate symmetry broken configurations.

We also compute the distribution of the time scale $\tau$, beyond which CSB occurs, $P(\tau)$ (Fig. 2.4(inset)). We find that $P(\tau) \sim \tau \exp(-\tau/\tau_0)$ with a mean $\langle \tau \rangle$ that scales linearly with $N$, for large $N$. Therefore, the mean-field result is recovered in the thermodynamic limit $N \to \infty$, when fluctuations are negligible.

This dynamical bifurcation is a generic consequence of fluctuations and shows up even in a simplified reaction scheme (Fig.2.1 (b)), with $k_1(R) = k_1N_R/N$ and $k_1(R^*) = k_1N_{R^*}/N$, where $N = N_A + N_R + N_{R^*}$. In this simple scheme: (i) the mean field equations do not show CSB, (ii) the chemical master equation shows CSB and (iii) the master equation goes over to the mean field equations in the limit $N \to \infty$.

We have thus demonstrated the role of chemical fluctuations in driving the autocatalytic system towards bimodality, when the mean-field has unique monostable solutions. CSB achieved is spontaneous (no initial bias), achieved exponentially fast, complete (100%) and arises purely due to the interplay of stochasticity and autocatalysis.
2.7 Cellular application(s)

Given the generality of these results, and the widespread prevalence of stochasticity in the environment of the cell, we expect that this phenomenon should be manifested in several cellular processes going beyond the scale of the molecule to networks of interacting molecules. For instance, binary outcomes are an integral part of cell signaling and gene regulatory networks (Section 1.2). Feedback loops, often mediated by enzymatic catalysis and autocatalysis are the underlying motifs in such networks, playing a key role leading to binary cellular decisions.

Here we consider the following two instances: (i) Downstream signaling in response to ligand binding, resulting in distinct states of the T-cell (ii) Genetic switch involved in λ-phage life cycle.

2.7.1 Selection of agonism/antagonism

The immune response of the T-cell receptors(TCRs) on the surface of T-cell, upon binding to the pathogenic peptides on the surface of antigen presenting cells (APC), begins with the phosphorylation of TCRs. They can recognise as few as three foreign pathogenic peptides (agonists) amidst thousands of endogenous proteins present on APC, rendering it a sensitive machinery. Such recognition turns on downstream signaling cascade resulting in T-cell activation (agonism) and immune response. However if the agonists on the APC are mutated (antagonists), its binding shuts down (antagonism) further signaling. How such a spontaneous selection of binary outcome (agonism/antagonism) is invoked in response to the type of ligand binding?

Reference (5) analyses membrane proximal signaling (Fig. 2.5) in T-cells, taking into account the competing feedback loops mediated by kinase (Lck) and phosphatase (shp), which result in a purely stochastically driven
Figure 2.5: Binary decision in T-cell signaling: Feedback network for membrane proximal T-cell signaling (5). The coarse-grained network shows the minimal model, where the positive feedback from kinase (E) leads to agonism while that from phosphatase (S) leads to antagonism (similar to ours). These two competing loops coupled with stochasticity is sufficient to achieve spontaneous selection of a binary outcome.

Binary decision (agonism/antagonism). The binding of T-cell receptor (TCR) and the agonist (pathogen’s peptide, $A_1$) results in Lck-mediated phosphorylation, which is maintained in an on-state (agonism) with positive feedback from downstream signaling product (ERK, E). On the other hand, the binding of the antagonist (mutated pathogen’s peptide, $A_2$) turns it off (antagonism), with negative feedback from the phosphatase (shp, S) which locks the system in the off-state. Using a set of model reactions, they show that *irreversibility, branching* and *feedback regulations* are the necessary and sufficient conditions for stochastic bimodality in the absence of deterministic bistabilities.

While the analysis of (5) is complete, we show how this is an instantiation of our very general result. This is best done by constructing a simpler coarse-
grained network (Fig. 2.5) from their more elaborate one. The resemblance of the coarse-grained network to our general autocatalytic model is immediate. Following our general analysis, we should expect that this network should exhibit the following features –

(a) the deterministic steady-state is monostable.

(b) the spontaneous selection of agonism or antagonism in T-cell, in response to ligand binding, is achievable only in the presence of stochasticity, as noted in (5).

### 2.7.2 Selection of Lysis/Lysogeny

![Image of binary decision in Bacteriophage λ](image)

Figure 2.6: Binary decision in Bacteriophage λ: (a) A proposed exclusive-or switch where the binding of transcription factors on overlapping promotor regions result in mutual repression (4) of gene A (red) and B (blue) (b) Our model where gene A (red) and gene B (blue) compete autocatalytically (without repression) for overlapping promotor regions.

Often a pair of interacting network of genes flip-flop between the stable binary states, which further translates into distinct states of a cell. One such network is presented by the gene regulatory system in bacteriophage lambda—a bacterial virus. Upon invading the E.coli cell, two gene regulatory proteins, namely the lambda repressor protein (cI) and Cro protein, synthesized by the virus ultimately let it remain dormant (lysogeny state) or kill the host cell (lysis state) respectively. Both of them share a overlapping promotor region in the viral genome and hence binding of one regulatory protein to the promotor region hinders the simultaneous binding of the other.
Ref (4) analyses the selection of a subset of states (bound/unbound, unbound/bound) in λ-phage regulation. The repressor proteins cl and Cro compete for overlapping promoter regions. They propose an exclusive-OR switch, where binding of one gene product mutually excludes the other (Fig. 2.6(a)), due to overlapping promoter regions, given by,

\[
\begin{align*}
\frac{\partial_t n_R}{\partial t} &= k_1 f_1 (n_R, n_{R^*}) - k_2 n_R \\
\frac{\partial_t n_{R^*}}{\partial t} &= k_1 f_2 (n_R, n_{R^*}) - k_2 n_{R^*}
\end{align*}
\]

(2.7) (2.8)

where the coupling term

\[
f_1(n_R, n_{R^*}) = \frac{g_1(1 + n_R) + g_0 n_{R^*}}{1 + n_R + n_{R^*}}
\]

and an identical form for \( f_2 \) accounting for repression of \( R^* \) by \( R \). Since binding of one repressor mutually excludes the binding of other, if a repressor binds successfully then its concentration increases, further enhancing its probability of binding over the other. This in turn stably maintains the other gene repressed.

These essential features can be captured by the simpler network (Fig. 2.6(b)). The repressor proteins compete for the promoter regions A & B—coupling their dynamics. Stochastic selection of a repressor leads to its autocatalytic multiplication, which translates into enhancement of its chances of binding over the other. Such a sequence of events result in stable repression of the other gene.

Thus this minimal model results in the spontaneous selection of lysis or lysogeny. Unlike (4), there is no need to invoke repression (no antagonism). Note that reversibility (with rate \( k_2 \neq 0 \)) is necessary to achieve complete
selection.

Thus it appears that the interplay between pure autocatalysis and stochasticity generically results in the spontaneous selection of a binary (or multivariate) outcome in chemical systems occurring over a wide spectrum of systems.