Chapter 13
Mechanisms of the Origin of Life

13.1 The Anderson Model of the Origin of Biological Information

Any theory purporting to account for life cannot avoid facing the fundamental question about how life originated on this planet. One of the most physically realistic models of the origin of biological information (and hence of life) that I know of was proposed by P. W. Anderson and his coworkers (1983, 1987) (see Fig. 13.1). The model was based on thermal cycling (i.e., the cyclical changes of the temperature on the surface of the earth due to its daily rotation around its axis) of an RNA "soup" presumed to be present somewhere on the primordial earth surface some 3.5 billion years ago. The following quotation from Anderson (1987) describes the key ideas behind his model:

> . . . The autocatalytic mechanism which must be at the core of any prebiotic evolution scheme is the complementary conjugation of polymeric molecules, nominally RNA. It is assumed that the thermal cycle periodically breaks up the weak conjugation bonds between RNA polymers, and at a later stage allows them to reconjugate randomly. Once two polymers have simultaneously conjugated with the same 'template', matching adjoining sequences (see the RNA double strands located on the bottom of Fig. 13.1; my addition), they are permitted with some probability to bond completely together, thus elongating the chain and reproducing a longer sequence of the 'template'. This is the basic autocatalytic process, while the basic energy source is a constant supply of energy rich monomers (or short sequences of 2 or 3 monomers) which are added at each cycle and can be joined to the sequences already present by the conjugation-thermal cycling process. To achieve realism and a reasonably steady state, we must also postulate an error probability and a probability of chain death and/or breaking.

Anderson based his model of the origin of biological information on the concept of "frustrations" imported from spin glass physics (van Hemmen 1983). *Frustrations* are observed in physical systems with three or more components, each being able to exist in at least two energy (or spin) states (conveniently designated as + and –, or up and down, with opposite signs attracting and identical ones repelling each other) but, no matter how their spins are arranged, there exists at

S. Ji, *Molecular Theory of the Living Cell: Concepts, Molecular Mechanisms, and Biomedical Applications*, DOI 10.1007/978-1-4614-2152-8_13,

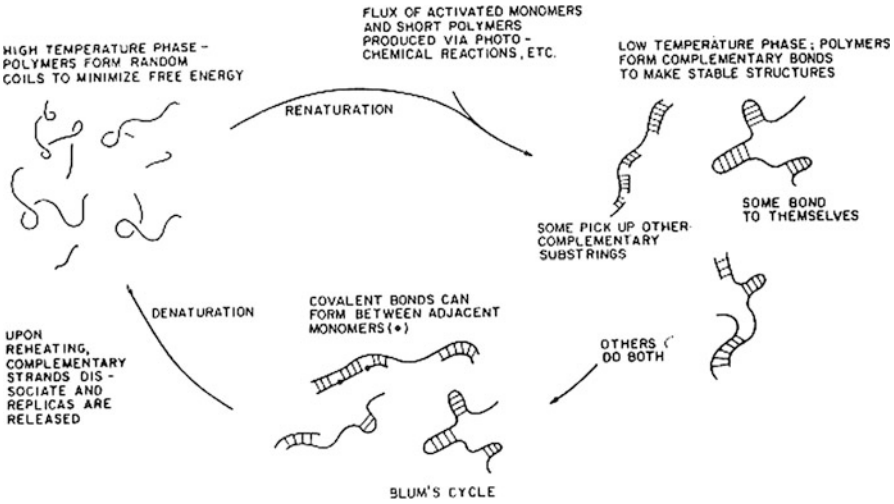

Fig. 13.1 The proposed model of the origin of biological information (and life) based on the concept of *frustrations* (Anderson 1983, 1987). Due to the presence of *frustrations*, some polymer chains cannot self-conjugate, thereby exposing single-stranded segments to environment to act as templates for self-replication (Reproduced from Anderson 1983, 1987)

least one pair of components whose spins are parallel to each other and hence of a non-minimal energy. Anderson and his colleagues represented the nucleotide sequence of an RNA molecule as a string of binary digits or spins, designating G as + +, C as − −, A as + −, and U as − + (which obeys the Watson–-Crick pairing rule). This allowed them to calculate the free energy (i.e., spin glass Hamiltonian, a mathematical function mapping spin configuration to the total energy of the spin system) of RNA molecules described as linear strings of spins. Furthermore, they defined what is referred to as the "death function" *D(S)* as a nonlinearly decreasing function of the *spin glass Hamiltonian*:

$$D(S) = 1/\{\exp[-H(S) + \rho N] + 1\} \tag{13.1}$$

where *H(S)* is the spin glass Hamiltonian (or the total energy of the spin system S), ρ is proportionality constant and *N* is the number of spins in the system (which is less than ten in the case studied in Fig. 13.2 below). Repeated applications of Eq. 13.1 to a collection of short RNA sequences showed that certain sequences died out with time (see the 7- and 8-mers in Fig. 13.2) whereas certain others (see the 11- and 12-mers) grew with repeated "thermal cycling," reminiscent of the selective growth of some nucleotide sequences in living systems. A similar finding was reported by Zeldovich et al. (2007a, b, 2008) (see Sect. 14.7).

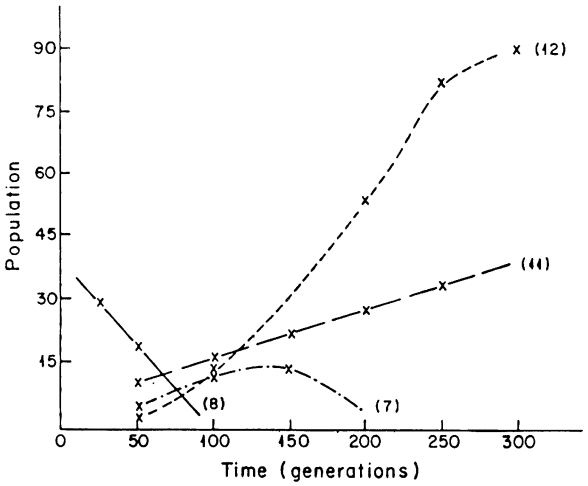

Fig. 13.2 The temporal evolution of RNA fragments obeying the death function, Eq. 13.1

13.2 The Conformon Model of the Origin of Life

Frustrations embedded in multicomponent physical systems including primitive RNA molecules embody both sequence *information* and mechanical *energy*, but Anderson and coworkers utilized only the *sequence information* in synthesizing complementary RNA fragments (see Fig. 13.1), thereby satisfying the *symbolic aspect* of Pattee's principle of *matter-symbol complementarity* (Pattee 1969, 1996; Ji 1999b) but not the energetic aspect. Consequently Anderson's model did not capitalize the mechanical (or conformational) energy associated with (or stored in) frustrations embedded in RNA to drive the synthesis of polymers. Anderson had to assume that "energy-rich" monomers, that is, nucleoside triphosphates (or nucleotides), were already available in the primordial RNA soup, but the presence of nucleoside triphosphates in the primordial soup may be very unlikely in view of its chemical instability, even if they were assumed to be formed by accidental coupling of five molecules belonging to the three different molecular classes – a base, a sugar, and an inorganic phosphate.

To overcome what I believed to be the deficiency of the Anderson model of the origin of biological information ("deficient" from the perspective of the *matter-symbol complementarity*), I modified the Anderson model by utilizing not only the *sequence information* (as he did) but also the *conformational energy* stored in frustrations (which he ignored). This is tantamount to assuming that the *frustrations* embedded in RNA molecules are *conformons* (sequence-specific conformational strains) (Ji 2000). The resulting *conformon-based* model of the origin of biological information (see Fig. 13.3) was named "the Princetonator" to reflect the facts (1) that it is an example of self-organizing chemical reaction–diffusion systems (as indicated by the suffix, *-ator*) and (2) that it is an extended version of the model of the origin of biological information originally proposed by Anderson and his

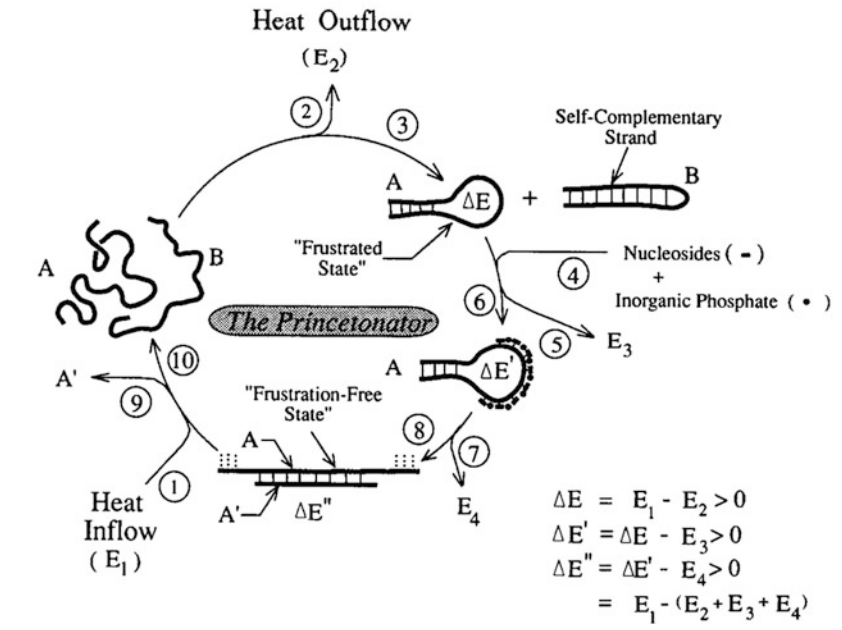

Fig. 13.3 A conformon-based model of the origin of self-replicating molecular systems that is constructed on the basis of the assumption that *frustrations* embedded in RNA carry both *sequence information* and *mechanical energy* and hence are examples (tokens, species) of *conformons* (Ji 2000) (Chap. 8). The key features of this model is that the thermal cycle of the earth's surface produce conformons in primitive RNA templates, which can drive the synthesis of RNA fragments that are complementary to a portion of the templates, the repetition of which leads to a complete replication of some RNA templates but not others. Conformons are equivalent to frustrations entrapped in sequence-specific loci in primordial biopolymers (Reproduced from Ji 1991)

group at Princeton (Fig. 13.2). The *Princetonator* contains the following key postulates (Ji 1991, pp. 224–225):

1. On the surface of the primordial earth about 3.5 billion years ago, there existed a pool (often called the "primordial soup") of at least two short biopolymers, A and B, most likely RNA molecules.
2. Due to thermal cycling (caused by the daily rotation of the earth or other cyclic motions on the earth such as tidal waves), the components of the primordial soup underwent periodic binding (e.g., due to low temperature; see Steps 3, 6, and 8 in Fig. 13.3) and de-binding (e.g., due to high temperature; see Step 10) processes.
3. During the low temperature phase, some biopolymers form a complete intramolecular binding (see B after Step 3) and some others form an incomplete intramolecular binding due to the presence of frustrations (see the bulge in A after Step 3) entrapping a part (ΔE) of the total energy flux, ($E_1 - E_2$), through

the primordial soup. The bulge (i.e., frustration) is located in sequence-specific sites and carries mechanical energy, ΔE, thus qualifying them as conformons (Ji 2000) (Sect. 8.1).
4. The bulge acts as a template for binding a set of monomers (i.e., nucleosides consisting of a ribose ring covalently linked to a base [symbolized as a dot connected to a bar] and inorganic phosphate ions [symbolized as a filled circle]) (Step 6).
5. The binding of the monomers and inorganic phosphate moieties to the bulge is postulated to trigger a conformational transition of the template causing covalent bond formation between nucleosides and adjacent inorganic phosphates to produce a string of nucleotides (see Step 8).
6. During the high temperature phase, the bound RNA fragments dissociate into monomers (see Step 10), producing unchanged B and A with a part of it reproduced (as A′), which has a finite probability of being elongated further through the repetition of the thermal cycling, eventually reproducing the original template A completely.

Pattee (1969) pointed out a set of logical and physical constraints that must be met by any satisfactory theory of the origin of life and biological information:

1. *The primeval ecosystem language* – The global set of geophysical and geochemical constraints of the primeval earth surface that were conducive to the spontaneous generation of self-replicating molecular systems or molecular *switches* (defined below).
2. *Complex molecular interactions leading to a very simple result – Communication* among molecules obey simple rules relative to the complex mechanisms underlying their *interactions*: Communication is in some way a simplification of complex dynamical interactions.
3. *Switches* – Physical devices whose function it is to turn on or off some physical or formal processes driven by energy dissipation. Networks of switches often referred to as "sequential switching machines" or "automata," can duplicate many of the most complex biological processes including human thought itself.
4. *Open-ended evolvability* – Not all self-replicating systems can also evolve. In order for self-replicating systems to evolve in an open-ended manner, special requirements additional to those of self-replication must be satisfied.
5. *Stability* – Of the many possible self-replicating systems that could have evolved spontaneously in the primeval ecosystem, only those with *stability*, *reliability,* and *persistence* survived.
6. The "von Neumann limit" – There exists a critical limit to the complexity of the network of switches which must be exceeded in order to effectuate self-replication. Since such a limit was first recognized by von Neumann, it is here suggested that the indicated limit be referred to as the "von Neumann limit," that is, the minimal level of the complexity (or organization) of the physical systems that are needed for an open-ended evolution.

All of these requirements appear to be satisfied by the combination of the original model of the origin of biological information proposed by Anderson (1983, 1987) and its modified version, the Princetonator (Ji 1991, pp. 224–225), as summarized in Table 13.1.

Table 13.1 The logical and physical requirements for the mechanism of the origin of *molecular messages* as specified by H. Pattee (1969) are met by the combination of the Anderson model of the origin of biological information (Fig. 13.2) and the Princetonator (Fig. 13.3)

Pattee's constraints on mechanisms of the origin of life (Pattee 1969)	Satisfied by	
	Anderson's model	The princetonator
1. Primeval ecosystem	The "RNA soup" on the surface of the earth subjected to thermal cycling	
2. Simple rules		*Conformon* production and utilization
3. Switches		Frustrated regions of RNA harboring *conformons* (see Fig. 13.3)
4. Open-ended evolvability		*Thermally accessible* conformations (called *virtual conformons* (Ji 1991, p. 136)) of RNA fragments that can drive self-replication when reified to real conformons upon coupling to exergonic chemical reactions, obeying the generalized Franck–Condon principle (Ji 1991, pp. 50–56) (Sect. 2.2.3). Due to thermal motions implicated, there is a finite probability of errors occurring during the conformon-driven copolymerization process, thus leading to mutations and open-ended evolution (Pattee 1995)
5. Stability		It is possible that n catalytically active molecular species (CAMS) must be colocalized in a small spatial volume (to be called the catalytic site) to effectuate spontaneous copolymerizations (see the Franck–Condon state defined on p. 433 in Ji 1974a and Fig. 1 in Ji 1979). If the probability of such a colocalization is P and the average probability of individual CAMS being located in the catalytic site is p, the following relation holds: $P = p^n$. This simple power law indicates that the *stability* and the *probability of spontaneous formation* of the self-replicating systems (SRS) increases and decreases, respectively, with increasing n. That is, the larger the value of n, the smaller is the probability P and the greater would be the stability of SRS against its accidental destruction by thermal motions
6. von Neumann limit		The *von Neumann limit* below which no SRSs can evolve may be identified with the exponent n in the relation, $P = p^n$, because n is determined by the balance between two opposing processes, namely, the spontaneous generation of SRSs (whose probability decreases with n) and the stability of SRSs (whose probability increases with n). We may refer to n as the *von Neumann exponent*

What is noteworthy about *the Princetonator* is that it is based on the same molecular entity, the conformon that has been found to account for the mechanisms of energy coupling underlying oxidative phosphorylation (Sect. 11.5), muscle contraction (Sect. 11.4.3), and active transport (Sect. 8.5) and of enzymic catalysis (Sect. 11.3.3). If the Princetonator is valid, it would mean that the same physical entity, the conformon that was thought to be responsible for the origin of life 3.5 billion years ago may be still responsible for the molecular mechanisms underlying life here and now. In the next section, it will be argued that the conformon also plays a crucial role in the evolution of life. These considerations motivate me to suggest the following postulate:

> The conformon not only was responsible for the origin of life 3.5 billion years ago but has also been responsible for the phylogeny and ontogeny of organisms ever since. (13.2)

We may refer to Statement 13.2 as the "Conformon Theory of the Origin of Life, Phylogeny, and Ontogeny (CTOLPO)." Since the origin of life, phylogeny, and ontogeny reflect aspects of life, Statement 13.2 could be alternatively called the "Conformon Theory of Life," which in turn is synonymous with the so-called *Fourth Law of Biology*, Statement 11.45, which simply states that

> Conformons are necessary and sufficient for life. (13.3)