The Chemoton: A Model for the Origin of Long RNA Templates.

Chrisantha Fernando and Ezequiel Di Paolo
Centre for Computational Neuroscience and Robotics.
Department of Informatics
University of Sussex
Brighton, BN1 9QH
{ctf20,ezequiel}@cogs.susx.ac.uk

Abstract

How could genomes have arisen? Two models based on Ganti's Chemoton are presented which demonstrate that under increasingly realistic assumptions, template replication is facilitated without the need of enzymes. It can do this because the template state is stoichiometrically coupled to the cell cycle. The first model demonstrates that under certain kinetic and environmental conditions there is an optimal template length, i.e. one which facilitates fastest replication of the Chemoton. This is in contradiction to previous findings by Csendes who claimed that longer templates allowed more rapid replication. In the second model, hydrogen bonding, phosphodiester bonding and template structure is modeled, so allowing dimer and oligomer formation, hydrolysis and elongation of templates. Here, monomer concentration oscillates throughout the cell cycle so that double strands form at low monomer concentrations and separate at high monomer concentrations. Therefore, this simulation provides evidence that a protocell with Chemoton organization is a plausible mechanism for the formation of long templates, a notorious problem for studies of the origin of life.

Introduction

The Chemoton, invented by Ganti (Ganti, 1971), and largely ignored since, is a model for a minimal protocell that provides a compelling hypothesis for the origin of long RNA template replicators, without the need for enzymes. Briefly, the problem of long template replication is that, beyond a certain length of double stranded RNA, double strands remain stuck together and cannot replicate. In modern cells, protein enzymes catalyze the separation of these strands. However, the length of the strand required to produce this enzyme is greater than the length of the strand that can spontaneously separate (Szathmáry, 2000). Therefore, some mechanism is required to facilitate strand separation before such protein enzymes came to be. This lead Eigen to propose the hypercycle, whereby RNA templates might act as enzymes to facilitate the replication of other RNA templates (Eigen, 1971). However, long templates were prevented from forming due to Eigen's paradox which resulted from the assumption that specific sequences of RNA template had different replication rates. This assumption does not hold in the Chemoton, because Ganti wanted to conceive of a minimal system that was capable of replication without any enzymes. In the simplest Chemoton it is only the length of template strands that effect replication rate. The plausibility of hypercycles as a solution to Eigen's paradox has been discussed elsewhere (Szathmáry and Maynard-Smith, 1997). Ganti proposed that the Chemoton could delicately control nucleotide concentrations, producing oscillatory conditions where at low levels, double strands formed, and at high concentrations, strand separation took place, so allowing template replication without enzymes. This paper demonstrates the mechanism of Ganti's proposal in simulation.

A review of the Chemoton is beyond the scope of this paper, and readers are referred to "The Principles of Life" (Ganti, 2003), "Chemoton Theory Vol I and II" (Ganti, 2004), and to Tibor Csendes' simulation of the Chemoton (Csendes, 1984). Briefly, the Chemoton is the simplest biologically plausible model of a primitive cell, which couples metabolism, membrane, and RNA templates, such that growth and division can occur under reasonable assumptions. In the simplest model, the sequence of templates is irrelevant, only the length carries information, i.e. regulates, the other two subsystems. Each subsystem is autocatalytic, i.e. regenerates itself, in strict coordination with the other two subsystems because the chemical reactions of each subsystem are coupled stoichiometrically. In the Chemoton, the high error rate of non-enzymatic template-directed template replication is of no concern, since it is only the length of templates, and later on in evolution, the ratio of nucleotides, that effect replication rate. This is because Ganti proposed that the original role of templates was merely as a "sink" to soak up waste products of metabolism, so regulating the rate of metabolism, and osmotic pressure. Thus, the Chemoton model provides an intermediate step in an evolutionary tale of the origin of genomes we know and love today. Just as feathers in birds might have initially been for warmth, and only later co-opted for flight, in the Chemoton, what seems to be the simplest possible coupling

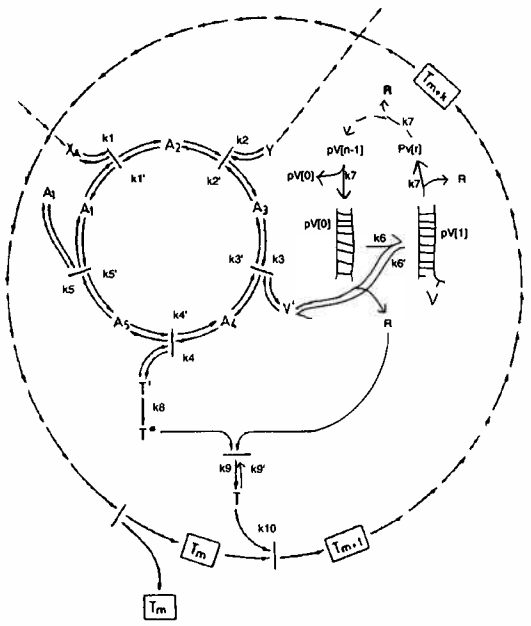

Figure 1: Adapted from Fig 4.6 of Chemoton Theory Vol I. The Chemoton consists of three stoichiometrically coupled autocatalytic chemical subsystems, a reversible metabolic subsystem based on the formose reaction, an irreversible membrane subsystem capable of growth and division, and an irreversible template polycondensation subsystem which serves as an informational control system potentially capable of unlimited heredity by a set of plausible evolutionary steps.

between genotype (template properties) and phenotype (metabolism and membrane) is proposed as the origin of long genomes.

Figure. 1 shows the chemical reactions of the Chemoton. Let us consider the metabolic subsystem first. A high energy precursor molecule, X, permeable to the membrane, T_m, reacts with metabolite, A_1, producing A_2. A_2 forms waste molecule Y (which diffuses out of the Chemoton), and another molecule A_3. A_3 produces V' , an impermeable template precursor monomer, and also produces A_4. A_4 produces T' , an impermeable membrane precursor molecule, and also produces A_5. A_5 produces two copies of A_1, thus completing the autocatalytic cycle. All members A_1 to A_5 are impermeable to the membrane. Metabolism is reversible, but the rate of reverse reaction is less than the forward reaction as free energy of the precursor X is greater than the free energy of the products Y, V', and T' of the cycle.

Next we consider the template subsystem. This system

replicates by template directed synthesis in which a preformed polynucleotide directs and catalyses the synthesis of its complementary form, from mono- or oligonucleotide building blocks. Monomer molecule V' comes together in a polycondensation reaction to form pVn (double stranded polymerised V' molecules of length n/2). This occurs only when V' is present in a concentration greater than the polycondensation threshold, $[V']^*$ [1]. So, $[V']$ increases until $[V']^*$ whereupon it begins to be bound to the polymer, also releasing molecule R, a hydrophilic component necessary for incorporation with T^* to form phospholipids, T, for the membrane. After N molecules of V' have bound to the double stranded polymer, the polymer splits up to produce two semi-conserved double stranded polymers, again of length n/2.

Finally, we consider the membrane subsystem. The T' molecule reacts irreversibly to produce T^* which binds to R to produce the phospholipid, T, which is spontaneously incorporated into the membrane. In each metabolic cycle, precisely equal quantities of T' and V' are produced. As the membrane grows, it has been shown experimentally that, if the volume does not increase rapidly enough to maintain the cell as a sphere, the cell divides into two daughter cells.

The three sub-systems regulate each other by feedback due to the existence of reversible reactions. When $[V']$ is less than $[V']^*$ (as long as no residue T^* and R are present) the volume does not increase because T cannot be formed. Thus, as the metabolic cycle functions, increasing $[V']$ hinders the operation of the metabolic cycle, i.e. the Chemoton slows down. The increasing concentration of intermediates creates an increase in the osmotic pressure within the Chemoton. Above $[V']^*$ the volume starts to increase as R is produced, but the number of osmotically active internal particles can never maintain the Chemoton as a sphere because V' is produced only in proportion to R. T is produced in proportion to surface area, at the same rate as V', and so V' cannot increase in proportion to the volume. Osmotic equilibrium can be reached only when the volume of the Chemoton has been reduced from that of a sphere. When the surface area of the sphere is doubled, the spherule divides into two spherules of approximately equal size by a well studied self-organizing process (Bachmann et al., 1992).

Model 1: Methods

The basic Chemoton described above is modeled using a set of standard kinetic equations with reaction rates the same as in Csendes' model and initial conditions which approximate the conditions at the start of cell division.

[1] $[X]$ means the concentration of X.

Metabolic sub-system

The metabolic subsystem is described by the same differential equations as used by Csendes. See Figure. 1 to see the meaning of the rate constants, k.

$$\frac{dA_1}{dt} = 2(k_5A_5 - k_5'A_1A_1) - k_1A_1X + k_1'A_2 \quad (1)$$

$$\frac{dA_2}{dt} = k_1A_1X - k_1'A_2 - k_2A_2 + k_2'A_3Y \quad (2)$$

$$\frac{dA_3}{dt} = k_2A_2 - k_2'A_3Y - k_3A_3 + k_3'A_4V' \quad (3)$$

$$\frac{dA_4}{dt} = k_3A_3 - k_3'A_4V' - k_4A_4 + k_4'A_5T' \quad (4)$$

$$\frac{dA_5}{dt} = k_4A_4 - k_4'A_5T' - k_5A_5 + k_5'A_1A_1 \quad (5)$$

Consider equation 1 for example. A_1 is produced by the reaction of an A_5 molecule to produce $2A_1$ molecules, hence the term $2k_5A_5$. Two A_1s are used up in the reverse of this reaction hense the term $2k_5'A_1A_1$. A_1s are also used up in reaction with X at rate k_1 and are produced in the reverse of this reaction at rate k_1'. By comparing the separate terms in each equation with the metabolic cycle in diagram 1, one can see they describe the stoichiometric equations correctly.

Template sub-system

We were unable to provide a meaningful physical interpretation for Csendes' model of template replication, in which he uses an "auxiliary variable" to represent template state. This implementation of template polycondensation is designed afresh, see Eqns. 6 to 10 and Figure. 1. The initial conditions consist of double stranded templates of length N/2 at concentration 0.01. Initiation is a reversible reaction taking place only at concentrations above [V']*. Propagation was an irreversible reaction, taking place also only when $[V'] > [V']^*$, on templates onto which at least one V monomer had already been bound. New full length double stranded templates,($pVn[0]$), were formed when a final V' monomer was bound to a strand of length $pV(n-1)$. Thus, we modeled the concentration of polymers at each discrete stage of replication, with a separate differential equation.

$$\frac{d(pVn[0])}{dt} = 2k_7pVn[N-1]V' + k_6'pVn[1]R - k_6pVn[0]V' \quad (6)$$

$$\frac{d(pVn[1])}{dt} = k_6pVn[0]V' - k_6'pVn[1]R - k_7pVn[1]V' \quad (7)$$

$$\frac{d(pVn[r])}{dt} = k_7pVn[r-1]V' - k_7pVn[r]V' \quad : n-1 \geq r \geq 2 \quad (8)$$

$$\frac{dV'}{dt} = k_3A_3 - k_3'A_4V' + k_6'pVn[1]R - k_6pVn[0]V' - \sum_{1}^{n-1} k_7pVn[r]V' \quad (9)$$

$$\frac{dR}{dt} = k_6pVn[0]V' - k_6'pVn[1]R + k_9'T - k_9T^* - k_9T^*R + \sum_{1}^{n-1} k_7pVn[r]V' \quad (10)$$

Membrane sub-system

The membrane precursor, T', is irreversibly converted into T^* which can then react with R (the by-product of polycondensation) in a reversible reaction to produce T. T then reacts in an irreversible reaction with the membrane, T_m, producing T_{m+1} in proportion to surface area S.

$$\frac{d(T')}{dt} = k_4A_4 - k_4'A_5T' - k_8T' \quad (11)$$

$$\frac{d(T^*)}{dt} = k_8T' - k_9T^*R + k_9'T \quad (12)$$

$$\frac{d(T)}{dt} = k_9T^*R - k_9'T - k_{10}TS \quad (13)$$

$$\frac{d(S)}{dt} = k_{10}TS \quad (14)$$

Calculation of Volume.

The calculation of volume, Q, is the least realistic calculation of the current model. It is simply calculated as $Q = S^{3/2}$, with an initial volume of 1.0. Thus, we assume that the Chemoton remains a sphere. Although this assumption is valid before $[V']^*$, after threshold it will not be accurate as the volume will be less than that expected of a sphere of that surface area, because the Chemoton has deformed due to decreasing internal osmotic pressure, thus causing water to leave the cytoplasm. At each time step the ratio, Q(t)/Q(t+1), is multiplied by each of the values found for Eq 1-13 above, thereby adjusting the concentrations of the substrates, so that if volume is decreasing, the concentration of the substrates will be increased.
A time-step of 0.0001 is used, and the fastest reaction rate constant is 100. This model is deterministic and continuous.

Model I: Results

The dynamics of model I are shown in Figure. 2. The concentration of metabolites, A_1 to A_5, increases as X is metabolized. $[V']^*$ is exceeded in the second half the cell cycle. After replication, residues of T , T^* , R, and T remain, and this results in slow growth immediately after cell division. Templates at all stages of replication are present at the point of cell division, and are passed onto the daughter cells. However, template replication does not continue immediately after cell division as $[V]^*$ is not yet reached. The Chemoton

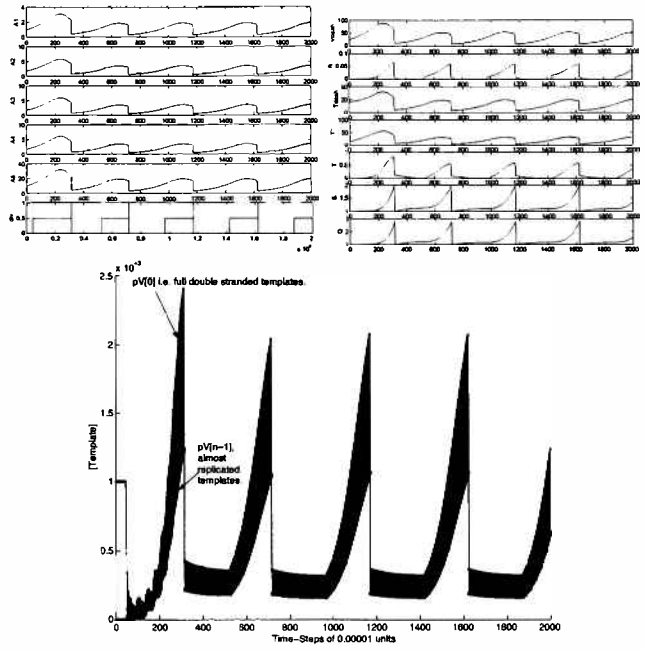

Figure 2: Top Left (1-5): Concentrations of A1-A5 over 4 cell cycles. Top Left (6): div = 0.5 when $[V'] > [V']^*$, and div = 1.0 at division. Top Right(1-6): V', R, T',T, Surface Area(S), and Volume (Q). Bottom: 25 lines, each showing the concentration of polymers (of length $N = 25$) at different stages of replication, i.e. $pV[0]$, is the full double stranded polymer, pV[1] has one extra V' bound, and $pV[n-1]$ requires only one more V' to be bound for separation into 2 semi-conserved $pV[0]$ polymers once again.

settles into a stabalized generation time, with a stable composition of molecules being passed to the progeny.

Many of Csendes' findings are confirmed. If the rate constants for the backward reactions of the metabolic cycle are made equal to the forward reactions, as would be the case if the free energy of X were not greater than the free energy of the products of the metabolic cycle, $[V']^*$ is not reached. Even if the Chemoton is initialized without V', T' and T^*, after a brief transient state a stable cell cycle is re-established, demonstrating the stability of the Chemoton to changes in initial state. Even if only A_1 is present, after a long transient a normal cycle is re-established. Csendes' findings that generation times are slightly longer for higher $[V']$ are confirmed.

If $[X]$ is reduced to 10, replication time increases from 0.455 to 0.65, and when $[X] = 1.0$ replication time is 2.4, that is slower by a factor of only 5.3 compared to the replication time at $[X] = 100$. The explanation given by Bekes (Bekes, 1975) for this non-linear adaptability of the Chemoton was that the concentration of $[A1]$ showed a compensating increase such that the product k_1XA_1 was maintained at a high value. A_1 buildup would be caused by the slowness

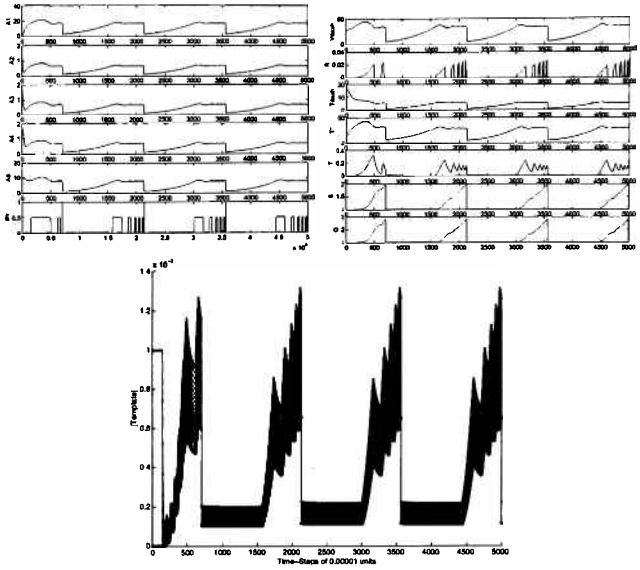

Figure 3: Chemoton functioning at $[X] = 1.0$, compared to previously where $[X]$ was 100.0. Since $[V']$ cannot be maintained above threshold permanently, (see top Left(6)),R is produced in spikes.

of the reaction incorporating X into the metabolic cycle, at low $[X]$. We indeed observed that when $[X] = 1.0$, $[A1]$ increased to a maximum of 20.0 from a previous maximum at $[X] = 100$ of 2.5. All other membrane chemicals had decreased in concentration. Also, at low $[X]$, the Chemoton functions by fits and starts because $[V']$ cannot be maintained above $[V']^*$ continuously (see Figure. 3).

Template Length and its Effect on Replication Rate.

Csendes claimed that Chemotons containing longer templates could more effectively compensate for decreases in nutrient concentration. If this is true, it would indicate a selection pressure for longer templates. However, using model I, for both high and low $[X]$ values, we found replication time increased with increasing template length, see Figure. 4. What accounts for these contradictory findings? Ganti writes, "A longer template molecule (i.e. larger N) requires more molecules V' for its reproduction, thus decreasing the $[V']$ and allowing the cycle to function more intensively. Conversely, a shorter template molecule allows the cycle to function more slowly and less intensively thus the template molecule pVn carries information concerning the system." So, what is happening to $[V']$ in our simulation? Examining the details of replication for a Chemoton with $N = 1000$ and $N = 3$ respectively, we see that $[V']$ responds in a manner opposite to that predicted by Ganti. At $N = 3$, $[V']$-max is apx. 35, for $N = 25$, $[V']$ max is apx. 85, and for $N = 1000$, $[V']$ max is apx. 1000. So, longer templates are 'mopping up' the V' molecules more rapidly, rather the

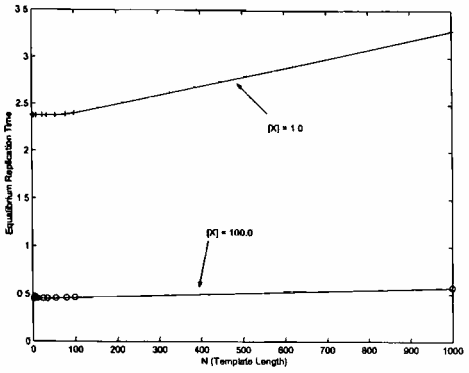

Figure 4: Replication time increases with longer templates, for both high and low nutrient concentration. Although difficult to see, the increase in replication time with increasing template length is monotonic.

opposite is the case.

In Csendes' model, and in ours so far, the rate of propagation of template polycondensation, ($k_7 = 10.0$) , is equal to the rate of initiation of polycondensation ($k_6 = 10.0$), therefore propogation proceeds at the same rate as initiation. According to these kinetics, it is merely the concentration of templates, irrespective of template length which will effect the rate of incorporation of $[V']$. What is more, with some consideration, it is predicted that template concentration will be less with increased template length, because a lower number of long templates are required to store all the V produced before cell division. This is indeed the case, for at $N = 3$, $pV[0]$ max = 0.4, at $N = 25$, $pV[0]$ max = 0.0025, and at $N = 1000$, $pV[0]$ max = 0.000007.

Are there any kinetic conditions for the basic Chemoton in which longer templates allow more rapid replication? What if $k_7 >> k_6$, such that the rate limiting step to polycondensation is initiation? We would expect that longer templates would have to undergo fewer rate limiting initiation steps before undergoing rapid proliferation, although, the same problem of lower template concentrations with longer templates would still persist. The same experiment as above was conducted but with $k_7 = 100.0$, see Figure. 5. The effect of template length in both conditions was very small. However, there was a qualitatively different profile compared with the previous experiment. At low $[X]$, templates of length 50 are optimal. This could be explained by a trade off between maintaining high template concentration and reducing the number of rate limiting initiation events.

Of-course these findings are crucially dependent upon our particular choice of template dynamics. Were it the case that longer templates provided a greater number of simultaniously accessible binding sites for proliferation, then the effect of reduced concentration could be counteracted. Alternatively, if the rate of proliferation of longer templates were greater, then the finding would also be reversed. The experiment above is conducted only on the most basic Chemoton without accounting for variable template length, enzymatic effects of specific template sequences upon metabolism, nor for the ability of templates composed of differing nucleotide ratios to simultaneously regulate alternative metabolic cycles. Thus, the relevance of such small effects of template length on replication rate may be of-course be irrelevant if other properties of long tamplates were to confer fitness.

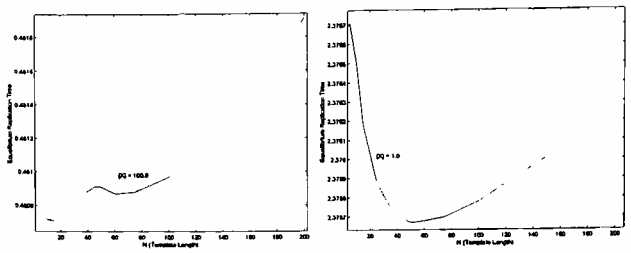

Figure 5: Left: At $[X] = 100.0$ the Chemoton exhibits slower replication with increased template length. The bump before this is unexplained. Right: At $[X] = 1.0$ the Chemoton shows a non-linear response to template length, with most rapid replication occuring at apx $N = 50$.

Model II: Methods

In order to see whether the Chemoton could still function with a slightly more realistic model of template replication, a simulation was written that was loosely based on models by Breivik (Breivik, 2001), Kanavarioti and Bernasconi (Kanavarioti and Bernasconi, 1990) and Wattis and Coveney (Wattis and Coveney, 1999). This model consists of monomeric units, V', interacting by both hydrogen (Watson-Crick type base pair bonds) and phosphodiester bonds. The probability of bond formation and breakage respond differently to environmental fluctuations in monomer concentration. Hydrogen bonds link strands to form double strands, and phosphodiester bonds link monomers lengthwise along a single strand. We assume that a phosphodiester bond is formed with very high probability, (0.5), in the configuration shown in Figure. 6. In all other configurations, the formation of hydrogen bonds is assumed to be more probable than phosphodiester bond formation. Also, we assumed that hydrogen bonds tend to break at fairly low monomer concentrations, whereas phosphodiester bonds are much more stable to high monomer concentrations.

The model is a discrete probabilistic model coupled to the continuous deterministic model described previously. The polymers are modeled as a variable size array of data structures shown in Figure. 6, and the probabilities of reactions

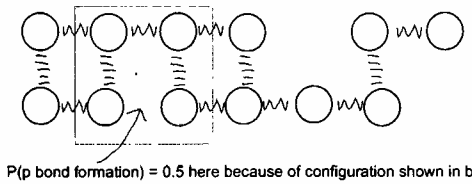

Figure 6: The data structure for a single polymer is a 2*M array, containing monomers, h-bonds and p-bonds. Potential bond sites can be deduced. The box shows the configuration necessary for high probability formation of a p bond, that is, the two monomers between which the p-bond can form must be hydrogen bonded to two monomers opposite. The two monomers opposite must already have a p-bond linking them to each other.

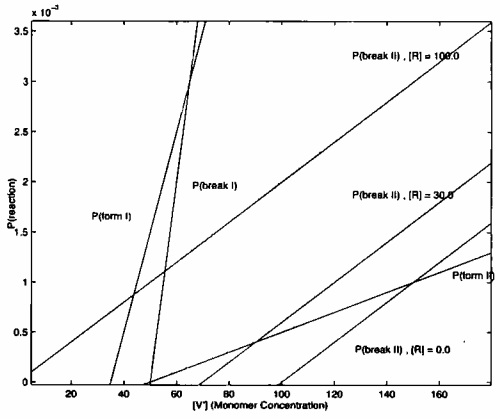

Figure 7: The dependency of the reaction probability of hydrogen and phosphodiester reactions upon $[V']$ and $[R]$ is shown. Bond I = hydrogen bond, and Bond II = phosphodiester bond. Where the lines for P(bond formation) and P(bond breakage) cross, there is an equilibrium.

are shown in Figure. 7.

For each site, the probability of reaction is $P(r) = 1 - e^{dE_r} : dE > 0$ and dE is calculated as shown in Eq. 15-19.

$$-dE(BondIform) = R_{hf}([V] - k_{hf}) \quad (15)$$
$$-dE(BondIbreak) = R_{hb}([V] - k_{hb}) \quad (16)$$
$$-dE(BondIIform) = R_{pf}([V] - k_{pf}) \quad (17)$$
$$-dE(BondIIbreak) = R_{pb}([R] + [V] - k_{pb}) \quad (18)$$

Where $R_{hf} = 0.0001, R_{hb} = 0.0002, k_{hf} = 35.0, k_{hb} = 50.0, R_{pf} = 0.00001, R_{pb} = 0.00002, k_{pf} = 50.0, k_{pb} = 100.0$. Thus, elongation and hydrolysis of chains can occur. However, several unrealistic features remain. We have not modeled the interactions between chains fully, i.e. reactions between polymers in separate data structures do not occur, so that spontaneous association of single-stranded oligomers are not modeled between data structures, only within data-structures. Detachment of oligomers is modeled, see below. The probabilities of binding and breakage,

except for the configuration in Figure. 6, are not dependent upon any structural feature of the polymer, e.g. a h-bond in the center of a double stranded polymer is equally likely to break as one on the edge of the polymer (i.e. no stacking reaction). Only one nucleotide, V' has been modeled. This reduces the problem of a combinatorial explosion in polymer configurations. Hydrolysis and dimerization of activated V' monomers is ignored, although dimers of V' can be formed by cleavage of longer strands. These decisions will clearly effect the findings, and future work must be to make these polymer dynamics more biologically plausible.

The simulation is initiated with 500 double stranded polymers of length 3. At each time-step, as well as calculating the membrane and metabolism equations, we calculate a probability of formation and breakage of type I & II bonds for each actual and potential binding site in each individual polymer. Based on these probabilities, the state of each polymer is changed. If two or more separate strands exist in the same data-structure, the detached strands (except one) are moved to new empty data-structures. Shorter oligomers dissociate faster than longer ones since more h-bonds need to be broken. The net change in $[V']$ caused by binding or release of V' from polymers is then calculated and incorporated into the differential equations for $[V']$ & $[R]$, after multiplication by a scaling factor (10). A sufficiently large number of polymers must be simulated to obtain a smooth net change in $[V']$. We assume R is produced in proportion to the incorporation of V' onto strands[2].

Model II: Results.

Figure. 8 shows the Chemoton initialized with 500 double stranded polymers of length 3, and with $[X]$ changed from 100 to 1.0 half way through the trial. The average polymer size drops from 6.0 to 5.5. The number of polymers (pV) is doubled rapidly when strands split, and this corresponds to a decrease in the average number of h-bonds. New p-bonds are formed immediately following the incorporation of monomers onto the single strand by the high probability p-bond formation when strands are in the configuration shown in Figure. 6. At low $[X]$, splitting of double strands by h-bond breakage is staggered since $[V']$ cannot be maintained above 50.0 for very long. Elongation takes place only rarely, since $[V']$ remains below 50.0 for much of the time.

[2] Future models will only have R produced when V' joins to the strand by a p-bond.

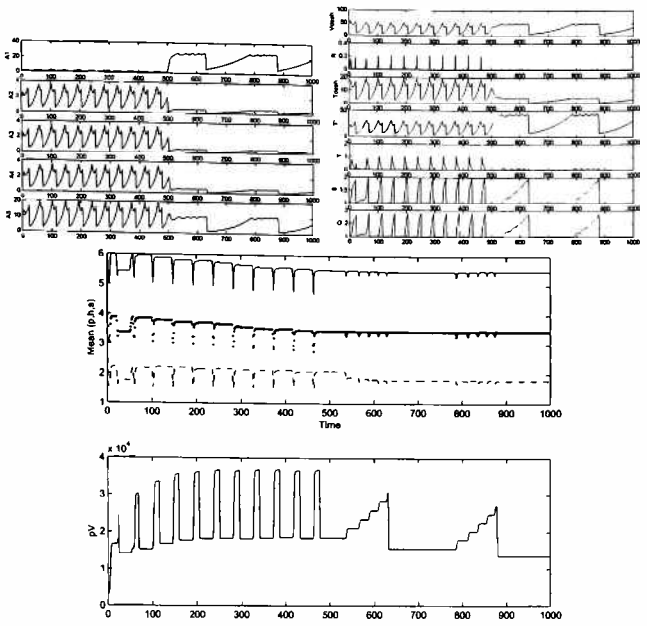

Figure 8: Model II: The response of Chemoton to reduction in $[X]$ is shown. Bottom(1): Mean (p,h,s) refers to mean number of p-bonds (dark line) , mean number of h-bonds (dashed line) and the mean polymer size (no. of monomers in a polymer). Bottom(2): pV refers to the total number of monomers in the genome.

Equilibrium template length depends on $[X]$ and p/h-bond probabilities.

Could changes in the relative probabilities of p-bond and h-bond formation effect the tendency for template elongation? The probability of p bond formation at low [V'] was increased by making $k_{pf} = 40.0$. Figure. 9 shows that, at high $[X]$ (100.0), elongation does occur until an equilibrium mean monomer number of 14 is reached, corresponding to double strands of length 7. At decreasing $[X]$ (1.0), there was a tendency for strands to elongate to longer lengths because $[V']$ did not exceed 50.0 as often, so that h-bonds were broken at a lower rate. This corresponds to a decrease in the total number of polymers being passed to each daughter cell.

Externally imposed $[V']$ oscillation verses Chemoton controlled $[V']$.

Figure. 10 (Top) shows the distribution of polymer sizes obtained in a trial with $k_{pf} = 35.0$ and $[X] = 50.0$. Compared to a trial in which oscillation of $[V']$ is imposed externally as a sine function, we see that a similar distribution of template sizes is obtained, see Figure. 10 (Bottom). However, to achieve this tight control of $[V']$ experimentally is extremely difficult, and such an experiment has not been conducted. The Chemoton provides a parsimonious explanation of how such oscillatory concentrations could be achieved.

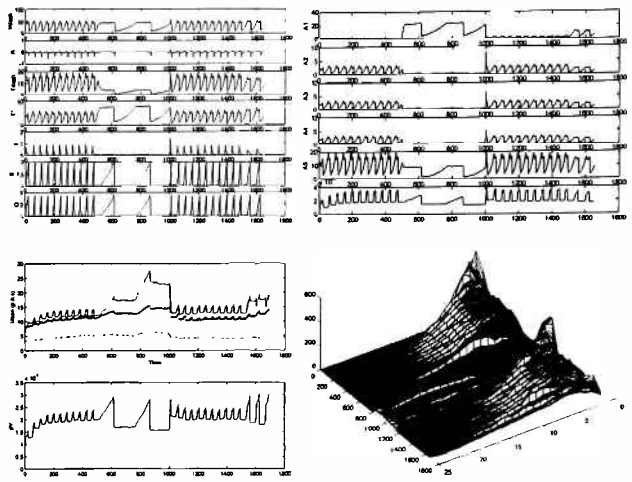

Figure 9: At First, [X] = 100.0, it is then decreased to 1.0, increased to 50.0 and reduced again to 10.0 $k_{pf} = 40.0$. Bottom Left: An equalibrium template length of 7 is reached with $[X] = 100.0$. When $[X]$ is reduced to 1.0, templates elongate further. Bottom Right: Absolute frequency of templates of different lengths, over the whole run.

Conclusions

In model I, it was demonstrated that under certain kinetic conditions, an optimal template length exists. Below this optimum length, template initiation is rate limiting, but above this length, although fewer initiation steps are required, the template concentration is decreased, so that propagation becomes rate limiting. In model II it was demonstrated that the equilibrium template length at any given $[X]$ was dependent upon the relative probabilities of h-bond and p-bond formation. When p-bonds could form at low $[V']$, at low $[X]$ there was a tendency for templates to grow as double strands and not separate as often, so decreasing the total template number passed onto each daughter cell. Thus, the length and concentration of templates conveyed information about the environmental conditions experienced by the parents. Finally, model II demonstrated that the tight coupling of the template state to the cell cycle in the Chemoton resulted in a different distribution of template sizes compared to that obtained when [V'] was oscillated by an external experimenter.

Further work will considerably improve the biological plausibility of the template dynamics, e.g. adding more realistic interactions between polymers in separate data-structures, adding the stacking reaction, and altering binding probabilities on the basis of experimental data. The hope is of providing a simple and viable alternative for experimentalists, (Rasmussen et al., 2003), of how an evolvable genotype-phenotype coupling could occur in a real protocell.

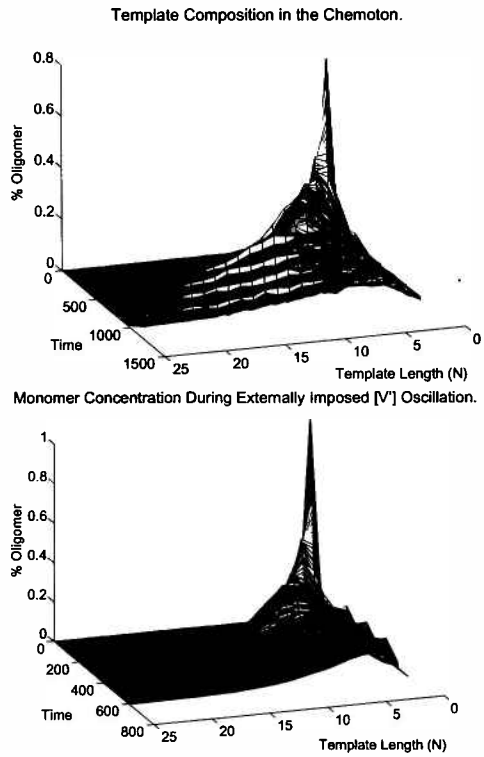

Figure 10: Top: Distribution of templates of average length N, calculated by summing the number of monomers in a polymer and dividing by 2, for the Chemoton, undergoing 4 divisions. Bottom: Distribution of templates for an in vitro study where $[V'] = 30 + 23 * \sin(500)(0.032)(0.00001).t$, is imposed externally to roughly match the $[V']$ observed in the Chemoton, but where $[R]$ is maintained at 0. $[V']$ is thus, un-reactive to template state. In practice, achieving this sine wave would be extraordinarily difficult. Slightly longer templates are obtained by the Chemoton.

Acknowledgements

Special thanks to Eors Száthmary for discussions, advice, references, and the idea for this paper. Also thanks to Simon McGregor, Inman Harvey, Phil Husbands, Naoaki Ono, and Tim Hutton. Thanks also to my parents and kind family in Sri-Lanka where this paper was written.

Appendix

Initial conditions in model 1 : [A1] = 1.0, [A2] = 1.8, [A3] = 1.9, [A4] = 1.7, [A5] = 10.0, [V'] = 26.0, [T'] = 17, [T*] = 14, [T] = 0.0, [R] = 0.0, [X] = 100.0, [Y] = 0.1, [pVn] = 0.01, Surface Area (S) = 1.0, Volume (Q) = 1.0, Polycondensation Threshold ([V']* = 35.0), Template Length (N) = 25.

Reaction Rate Constants in model 1: k1 = 2.0, k2 = 100.0, k3 = 100.0, k4 = 100.0, k5 = 10.0, k6 (if $[V'] > [V']*$) = 10.0 else k6 = 0, k7 = 10.0, k8 = 10.0, k9 = 10.0, k10 = 10.0 k1' = k2' = k3' = k4' = k5' = 0.1, k6' = 1.0, k9' = 0.1

References

Bachmann, P., Luisi, P., and Lang, J. (1992). Autocatalytic self-replicating micelles as models for prebiotic structures. *Nature*, 357:57–59.

Bekes, F. (1975). Simulation of kinetics of proliferating chemical systems. *Biosystems*, 7:189–195.

Breivik, J. (2001). Self-organization of template-replicating polymers and the spontaneous rise of genetic information. *Entropy*, 3:273–279.

Csendes, T. (1984). A simulation study of the chemoton. *Kybernetes*, 13:79–85.

Eigen, M. (1971). Self-organisation of matter and the evolution of biological macromolecules. *Naturwissenshaften*, 58:465–523.

Ganti, T. (1971). *The Principle of Life(In Hungarian)*. Budapest: Gondolat.

Ganti, T. (2003). *The Principles of Life*. Oxford University Press.

Ganti, T. (2004). *Chemoton Theory Vol I and II*. Kluwer.

Kanavarioti, A. and Bernasconi, B. (1990). Computer simulation in template-directed oligonucleotide synthesis. *Journal of Molecular Evolution*, 31:470–477.

Rasmussen, S., Chen, L., Nilsson, N., and Abe, S. (2003). Bridging living and non-living matter. *Artificial Life*, 9:269–316.

Szathmáry, E. (2000). The evolution of replicators. *Philosophical Transactions of the Royal Society of London. B.*, 355:1669–1676.

Szathmáry, E. and Maynard-Smith, J. (1997). From replicators to reproducers: the first major transitions leading to life. *Journal of Theoretical Biology*, 187:555–571.

Wattis, J. and Coveney, P. (1999). The origin of the RNA world: A kinetic model. *Journal of Physical Chemistry B.*, 103:4231–4250.