Non-Template Molecules Designed For Open-Ended Evolution

Chris Gordon-Smith[1]

[1]SimSoup
c.gordonsmith@gmail.com

Abstract

Theories of the Origin of Life can be categorised as 'template replication first' and 'metabolism first'. A key question for metabolism first theories is whether metabolic systems can support open-ended evolution; this is related to the number of possible persistent states of such a system. Earlier work[1] has demonstrated that artificial chemical systems can have memory; an essential requirement for inheritance. The current paper extends this, taking a 'proof of concept' approach to the question of the number of persistent states. It shows an artificial chemical network forming a 'memory bank' with many possible states. It also makes the link between chemical network structure and molecular structure, and provides a design for a set of artificial molecular species for the memory bank network. Preliminary simulation results from the SimSoup artificial chemistry simulator are included, confirming the operation of an initial set of 'memory units'. The work supports the view that open-ended evolution can begin without requiring highly complex template molecules.

Motivation, Approach, And Paper Overview

Metabolic theories of the Origin of Life propose that early organisms were metabolic systems that transmitted inherited information without the use of template replicating molecules such as DNA and RNA, and without the very complex mechanisms needed for their accurate replication[2].

It is envisaged that the systems were individuals capable of growth and reproduction; in some theories they are thought of as protocells that could divide. Variations in the metabolisms of different individuals would have led to differences in fitness that would drive evolution.

For this to be workable, successful variations would have to be retained and passed on to offspring. In addition, for evolution to be effective it would need to be open-ended, with a large number of possible variations in metabolism.

The motivation for this paper is to investigate whether this is feasible. A 'proof of concept' approach is adopted in which an artificial chemical network and associated molecular structures are *designed* for open-ended evolution. If the structures identified are not too complex, then it is reasonable to suppose that molecules with similar capabilities and properties could have occurred in the prebiotic world.

The rest of this paper includes the following:[3]

- *Conceptual Background* inspiring this work
- *Memory In Chemical Networks:*
  - *A Network Oriented View Of Chemistry:* A description of the *Network Components* from which chemical networks are constructed, the way these can be combined to form more complex *Compound Interactions*, an explanation of the distinction between *Static and Dynamic Networks*, and a discussion of *Catalysis* from a network point of view
  - *Network Memory And Exploration:* A description of a network that forms a *Two State Memory Unit*, and a discussion of how such units can be put together such that *The Dynamic Network Explores The Static Network*
- *Network Structure For High Memory Capacity:* A description of a network in which many *Memory Unit Sub-Networks* are combined to form a *Memory Bank Network*
- *Molecular Structure For The Memory Units:* A detailed description of a set of molecular structures in the SimSoup artificial chemistry that have been designed to implement the Memory Bank Network. This section describes:
  - *Molecular Structure In SimSoup*
  - *Atom Types For The Memory Unit Molecules*
  - The *Memory Unit Molecule Structures And Dimer Splitting:* Structures of Memory Unit Molecules, and of Dimer Splitting that is key to its operation
- *Results* of 'proof of concept' tests to investigate the workability of the memory bank
- *Discussion, Conclusions And Prospects, and References*.

[1]See Gordon-Smith (2005, 2007, 2009a,b) for earlier papers including SimSoup model details, and SimSoup (2011) for open source program code.

[2]Such mechanisms are prebiotically implausible, and so problematic for template replication first theories.

[3]Section names are italicised.

Conceptual Background

The SimSoup project takes inspiration from:-

- Metabolic theories including those of Aleksandr Oparin (Oparin, 1957), Stuart Kauffman (Kauffman, 1993), Freeman Dyson (Dyson, 1999), Chrisantha Fernando and Jonathan Rowe (Fernando and Rowe, 2007), and the Lipid World theory and GARD model of Doron Lancet's group (Segré et al., 1998, 2001a,b)
- Graham Cairns-Smith's clay crystal and genetic takeover theory (Cairns-Smith, 1982)
- Tibor Gánti's work on the principles of life and chemoton theory (Gánti, 2003)
- Network theory, particularly the work of Sanjay Jain and Sandeep Krishna (Jain and Krishna, 1998; Krishna, 2003)
- The Chemical Organisation Theory of Peter Dittrich and Pietro Speroni di Fenzio (Dittrich and di Fenizio, 2007)
- Günter Wächtershäuser's chemo-autotrophic Iron-Sulphur World (Wächtershäuser, 1990, 1997, 2006)
- Linus Pauling's chemical bond theory (Pauling, 1960).

Memory In Chemical Networks

A Network Oriented View Of Chemistry

This section presents a network oriented view of chemistry, and introduces terminology used in SimSoup.

Network Components The basic units of chemistry are particles and elementary reactions between these particles. The particles can be molecules or ions and are of different types (species). In an elementary reaction, one or more particles reacts directly to form products in a single reaction step and with a single transition state.

In SimSoup, a species of particle is called a *Molecule Type*, and an elementary reaction with particular Reactant(s) and Product(s) is called an *Interaction Type*. An instance of a Molecule Type is a *Molecule*, and an instance of an Interaction Type is an *Interaction*.

From a network point of view[4], there are only three forms of elementary reaction as follows (see Figure 1):-

- *Construction*: Two Reactant Molecules join to form a single Product Molecule
- *Transformation*: A single Reactant Molecule re-arranges to form a Product with the same atomic composition, but different structure
- *Fission*: A single Reactant Molecule splits to form two Product Molecules.

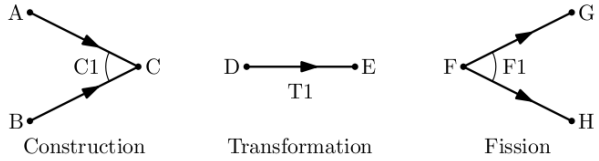


Figure 1: The three forms of Interaction Type. In Construction C1, Reactant Molecules of types A and B join to form a Product of type C. In Transformation T1, a Molecule of type D re-arranges to form a Molecule of type E. In Fission F1, a Molecule of type F splits into Molecules of types G and H.

Compound Interactions More complex reactions can take place as a result of Interaction Types combining in various sequences. Figure 2 shows a compound interaction with overall scheme A + B $\rightarrow$ E + F.

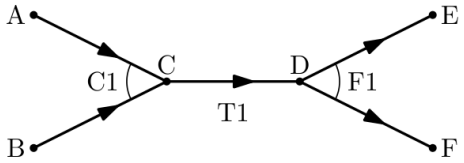


Figure 2: A Compound Interaction

A Compound Interaction does not have a rate constant that determines the reaction rate according to the concentration(s) of the (non-intermediate) Reactant(s). In Figure 2, the reaction dynamics depend on the concentrations of C and D, as well as of A and B. If the Compound Interaction forms part of a larger network, C and D may be Reactants or Products for other Interaction Types, and so the reaction depends on factors other than the concentrations of A and B.

Static And Dynamic Chemical Networks A set of Molecule Types and Interaction Types (along with temperature and pressure dependent rate constants) define a *static network*. This is determined for all time by the laws of physics. A *dynamic network* is a set of actual Molecules and actual Interactions taking place between them at particular rates. As such, it is a possible process that can occur within the framework of a static network.

[4] A network constructed from elements as shown in Figure 1 is *not* a graph in which the vertices represent Molecule Types and the edges represent Interaction Types. Constructions and Fissions each have three vertices connected by two edges, whereas each edge in a graph has only two vertices.

A chemical network can be represented by a *directed bipartite graph*. A bipartite graph has vertices that can be divided into two disjoint sets U and V such that every edge connects a vertex in U to one in V. In a directed bipartite graph, each edge has a direction.

Alternatively, a chemical network can be represented by a *directed hypergraph*; a hypergraph is a generalisation of a graph in which a 'hyperedge' can connect any number of vertices. In a directed hypergraph, hyperedges connect 'head' vertices to 'tail' vertices. The network elements of Figure 1 can be regarded as edges in a directed hypergraph.

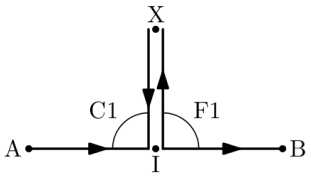


Figure 3: Catalysis Example

Catalysis The word 'Catalyst' does not denote a kind of Molecule. It denotes a *role* that a Molecule can play in a chemical process. In Figure 3, X plays the role of a catalyst; it is used by Construction C1, and released by Fission F1, so that overall it is neither consumed nor produced.

Network Memory And Exploration

A Two State Memory Unit Figure 4 shows a simple (static) network for an artificial chemistry consisting of three elementary reactions C1, F1 and F2:

$A + X \rightarrow I1$	(Construction C1)
$I1 \rightarrow I2 + X$	(Fission F1)
$I2 \rightarrow B + X$	(Fission F2).

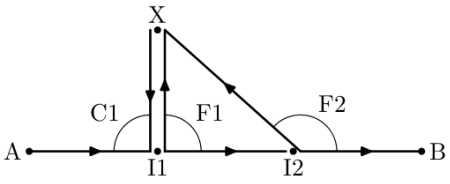


Figure 4: A two state chemical memory unit.

A is abundantly available 'food'; initially no other Molecules are present. In the absence of X Molecules, Construction C1 cannot proceed and A remains the only Molecule Type present. If a single Molecule of X is introduced, a Molecule of I1 is produced (Construction C1). This then splits (Fission F1) to release an X Molecule and an I2 Molecule. The I2 Molecule then splits (Fission F2) to release another X Molecule plus a B Molecule. Overall, for each A Molecule consumed, one new X Molecule becomes available in addition to the B Molecule. As a result, the supply of X is maintained (even if there is some 'leakage').

The network is bistable; it has two states, one in which only A Molecules are present and no Interactions occur, and another in which Interactions proceed and X is maintained. The introduction of a single Molecule of X is 'remembered' because it triggers a switch to a new persistent state.

The network therefore constitutes a simple memory unit with an information capacity of 1 bit[5].

[5]Under the current design for the memory unit, state changes of the unit are not reversible. However, such changes *can* be reversed at the ecosystem level. See the 'Discussion' section below.

The Dynamic Network Explores The Static Network Figure 5 shows a static network in which two of the memory units in Figure 4 are connected in series.

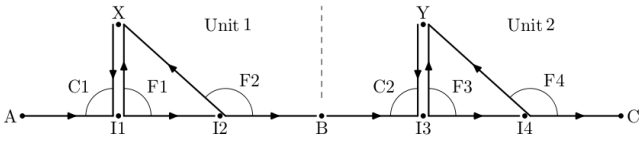


Figure 5: A two unit memory network with three states

If only A is available as 'food', there are three possible persistent states of the dynamic network: i) neither unit is active (only A is present), ii) only unit 1 is active, iii) both units are active.

In a more general situation where the static network is (effectively) infinite, we can consider a dynamic network to be 'exploring' the static network. A perturbation (such as the addition of a single X or Y molecule) can cause new parts of the network to become accessible.

Network Structure For High Memory Capacity

The previous section described how simple two-state memory units can be combined to form a larger network with more stable states and so higher memory capacity.

This section presents a network that systematically combines a large number of memory units to form a network with a correspondingly large memory capacity.

Memory Unit Sub-Network

Figure 6 shows a two state network that will form a memory unit within a larger 'memory bank' network[6].

$\mathrm{P_{s,p-1}}$ and $\mathrm{D_{sp}}$ are 'food'. The Interaction Types in the network are as follows:

$\mathrm{P_{s,p-1}} + \mathrm{M_{sp}} \rightarrow \mathrm{P_{sp}}$	C1 (Construction)
$\mathrm{D_{sp}} + \mathrm{P_{sp}} \rightarrow \mathrm{P_{sp}D_{sp}}$	C2 (Construction)
$\mathrm{P_{sp}D_{sp}} \rightarrow \mathrm{P_{sp}M_{sp}} + \mathrm{M_{sp}}$	F3 (Fission)
$\mathrm{P_{sp}M_{sp}} \rightarrow \mathrm{P_{sp}} + \mathrm{M_{sp}}$	F4 (Fission)

If a Molecule of $\mathrm{M_{sp}}$ is added to the food, then an Interaction of each of the four types can take place in sequence (C1, C2, F3, F4). The overall scheme for this sequence of Interactions is:

$$\mathrm{P_{s,p-1}} + \mathrm{D_{sp}} + \mathrm{M_{sp}} \rightarrow \mathrm{P_{sp}} + 2\mathrm{M_{sp}}$$

The sequence can only proceed if at least one Molecule of $\mathrm{M_{sp}}$ is present, but once the reaction has started it continues due to the excess production of $\mathrm{M_{sp}}$.

There is nothing 'special' about the sequence C1, C2, F3, F4. If the Interactions are considered in different sequences

[6]Molecule Type name convention: $\mathrm{M_{sp}}$, $\mathrm{P_{sp}}$ and $\mathrm{D_{sp}}$ indicate a monomer, polymer and 'closed dimer' respectively. See the section covering molecular structure for further explanation.

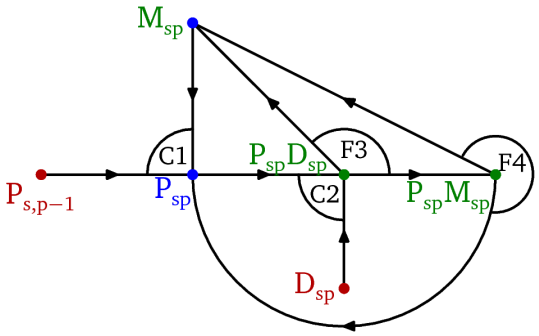


Figure 6: A memory unit for the 'memory bank'. Molecule Types are colour coded as follows: red - input to this unit, blue - output from this unit to the next unit, green - intermediate products. The blue and green Molecule Types together form an autocatalytic set. If the inputs are present, a Molecule of any member of the set can activate the network.

then it can be seen that a single Molecule of any one of $\mathrm{P_{sp}}$, $\mathrm{P_{sp}D_{sp}}$ or $\mathrm{P_{sp}M_{sp}}$ (in addition to the food) is also sufficient to activate the network.

In short $\mathrm{M_{sp}}$, $\mathrm{P_{sp}}$, $\mathrm{P_{sp}D_{sp}}$ and $\mathrm{P_{sp}M_{sp}}$ are an autocatalytic set that can be activated by any member of the set.

A Memory Bank Network

Figure 7 shows a 'memory bank' of 25 units in five independent rows or *series*. Each units has a label $\mathrm{U_{sp}}$, where s indicates the series, and p indicates the *position* of the unit in its series. Each unit has the structure shown in Figure 6, with only the specific Molecule Types varying. The large circles on the left of the diagram represent a maintained food set.

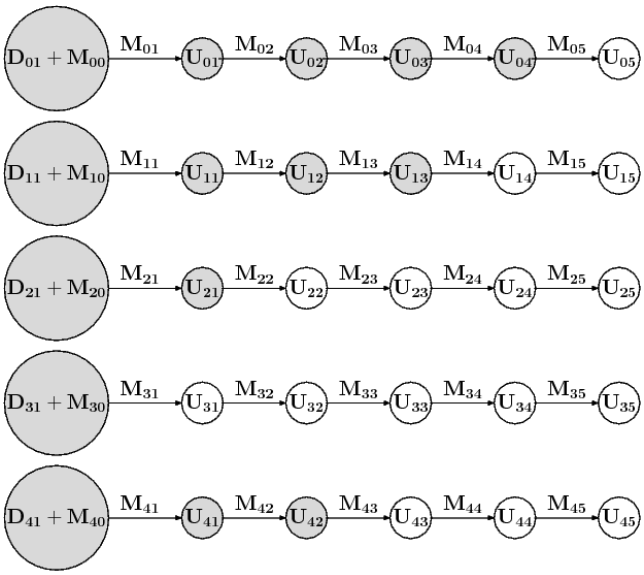


Figure 7: A Memory Bank with 25 units.

In each series the food provides the input to the first unit, and the outputs of each unit provide the inputs to the next unit. Each unit in a series may be either active or inactive; shading indicates an active unit. The next unit in a series can only become active if its predecessor is active (the maintained food set is considered to be the predecessor of the first unit, and is always active). The labels of the form $\mathrm{M_{sp}}$ over the arrows represent Molecule Types that will, if introduced in very small quantities, activate unit $\mathrm{U_{sp}}$ provided its predecessor is active.

Overall, the diagram represents a static network in which each of the five series has 6 possible states (from no units active, to all five active), so that the network as a whole can have $6^5 = 7776$ different states. A network with ten series of nine units would have 10^{10} possible states.

Molecular Structure For The Memory Units

In this section the link between network structure and molecular structure is made. A set of SimSoup Molecule Types that produce the memory bank of the previous section is described.

Molecular Structure in SimSoup

The approach to modelling molecular structure has been described elsewhere (Gordon-Smith, 2009b). It is summarised here, and an extension introduced for the work discussed here is described.

Molecules are two dimensional rigid structures built from Atoms bonded together such that they occupy fixed positions on a square 'Board' (similar to a chess board). Each square contains at most one Atom. Bond angles are always 90° or 180°, and bond lengths are all equal. Atoms bond together in a way broadly consistent with valence bond theory.

Molecules can Join or Split to form Molecules of different types. Joining must respect the 'one Atom per square' rule. Splitting occurs by breaking the weakest set of bonds that hold the Molecule into a single unit.

Bond strengths are usually fixed according to the types of Atom at each end of a bond. The extension introduced for this work is that in some cases, a bond can be perturbed (weakened or strengthened) by the proximity of Atoms that do not themselves participate in the bond.

Atom Types For The Memory Unit Molecules

The SimSoup Atom Types used for the Memory Unit Molecule Types are described below:

- Assemblite: Forms two bonds. Can be used to assemble the structural framework for a Molecule. Colour: black
- Stoppite: Forms one bond, and when present at a bonding site stops further growth of the Molecule at that site (much as Hydrogen does in an organic molecule). Colour: grey

- Junctium: Forms three bonds. Can be used to provide a 3 way junction in a structure. Colour: blue
- Loosium[7]: Forms three bonds. Can provide a weak (loose) bonding site within a structure. Does not bond to Anti-Loosium. Colour: spring green
- Anti-Loosium: Forms three bonds. Can provide a weak (loose) bonding site within a structure. Does not bond to Loosium. Colour: cyan
- Grabite: Forms three bonds. Can provide a bonding site in one monomer for another monomer to 'grab' as part of building a polymer. Colour: red
- Hookite: Forms two bonds. Can provide a 'hook' that can attach to an atom of Grabite to form a bond as part of building a polymer. Colour: green
- Perturbium: Forms three bonds. Bonds can be weakened or strengthened by nearby Metal atoms. Colour: magenta
- Metal: Forms one bond. Can perturb nearby Perturbium/Perturbium bonds, even though not bonded to Perturbium. Colour: orange.

Memory Unit Molecule Structures And Dimer Splitting

Monomers, Polymers And Closed Dimers This section describes the structures of Molecule Types that appear as Reactants for Constructions C1 and C2 in Figure 6.

Molecule Types of the form $\mathrm{M_{sp}}$ are *monomers*, those of the form $\mathrm{P_{sp}}$ and $\mathrm{P_{s,p-1}}$ are (short) *polymers*, and those of the form $\mathrm{D_{sp}}$ are *closed dimers*.

Figure 8 shows examples. Figure 8a shows monomer $\mathrm{M_{01}}$ and its structural units. The positions of the two recesses labelled $\mathrm{S} = 0$ and $\mathrm{P} = 1$ vary as the series s and position p indices vary. The recesses are called the *series recess* and the *position recess* respectively.

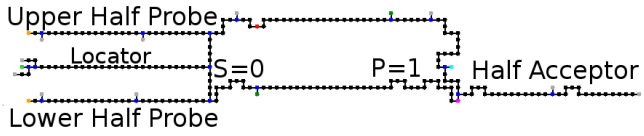
Along the top of each monomer are three small projections and a recess. The left hand *series projection* is directly above the series recess. The middle *position projection* is one place to the left of the position recess.

Figure 8b shows polymer $\mathrm{P_{01}}$. The naming convention for polymers is such that $\mathrm{P_{sp}}$ represents a polymer of length $p + 1$ whose end monomers are $\mathrm{M_{s0}}$ and $\mathrm{M_{s,p}}$.

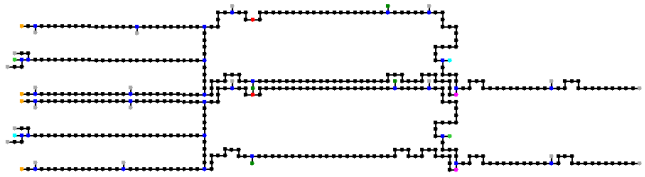
The positions of the recesses and projections on the top and bottom of the monomers ensure that two monomers can only join in a polymer if they are in the same series (same s index) and their position (p) indices differ by 1.

The Half-Probes and Half-Acceptor on each monomer also have recesses/projections, and the positions of these are similarly dependent on the series and position indices.

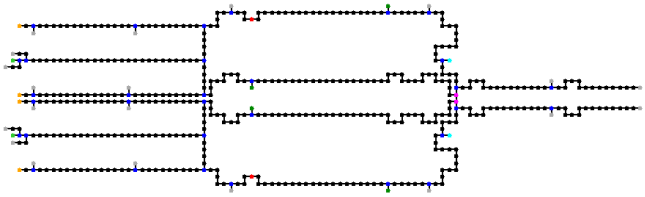
[7]Only two of the bonds supported by Loosium and Anti-Loosium are used for the memory unit Molecules.



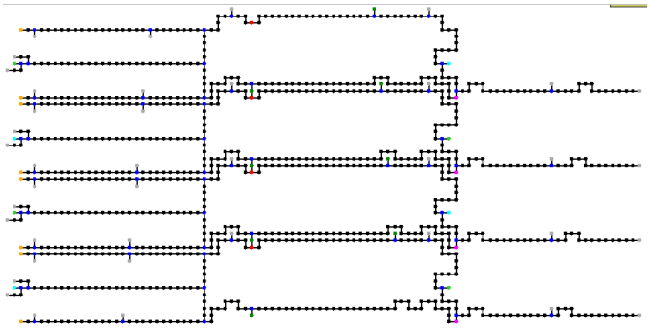
(a) Monomer $\mathrm{M_{01}}$ showing structural units



(b) Polymer $\mathrm{P_{01}}$



(c) Closed Dimer $\mathrm{D_{01}}$



(d) Polymer $\mathrm{P_{03}}$

Figure 8: Example monomer, closed dimer, and polymers. See the supplementary material for larger examples.

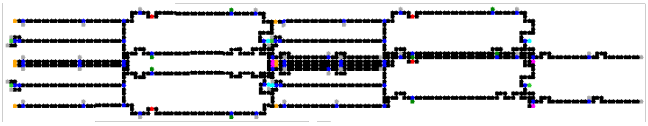
The structure of the monomers allows for both s and p to vary between 0 and 9. There are therefore 100 possible monomer types, and these can be used to construct 10 series of polymers, with polymers in each series being built from up to 10 monomers. Each series corresponds to a row in an enlarged version of the memory bank of Figure 7.

Figure 8c shows a closed dimer, formed by joining two monomers 'back to back'.

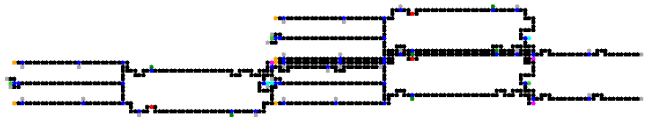
Finally, Figure 8d shows P_{03}, a polymer of length 4.

Dimer Splitting Intermediates And The Splitting Mechanism Figure 9 shows the structure of the Fission Reactants in Figure 6. Figure 9a shows Molecule Type $\mathrm{P_{01}D_{01}}$. Figure 9b shows Molecule Type $\mathrm{P_{01}M_{01}}$.

Dimer splitting is a key mechanism for the memory unit. It provides the means by which the autocatalytic set of Figure 6 maintains itself. Taking the example of Figure 9,



(a) Memory Unit intermediate $\mathrm{P_{01}D_{01}}$



(b) Memory Unit intermediate $\mathrm{P_{01}M_{01}}$

Figure 9: Dimer splitting intermediates

a Molecule of $\mathrm{P_{01}D_{01}}$ first splits (Fission F3) to release a Molecule of $\mathrm{P_{01}M_{01}}$ plus an $\mathrm{M_{01}}$ monomer, and then the $\mathrm{P_{01}M_{01}}$ Molecule splits (Fission F4) to release a $\mathrm{P_{01}}$ Molecule plus a second $\mathrm{M_{01}}$ monomer.

In short, the autocatalytic set maintains itself by splitting a 'food' dimer $\mathrm{D_{sp}}$ to produce a surplus of the monomer $\mathrm{M_{sp}}$.

Dimer splitting involves a mechanism in which a polymer temporarily binds a dimer, and as result the Perturbium/Perturbium bond that holds the dimer together is weakened. The details of this can be explained by reference to Figure 10, which shows the central part of $\mathrm{P_{01}D_{01}}$.

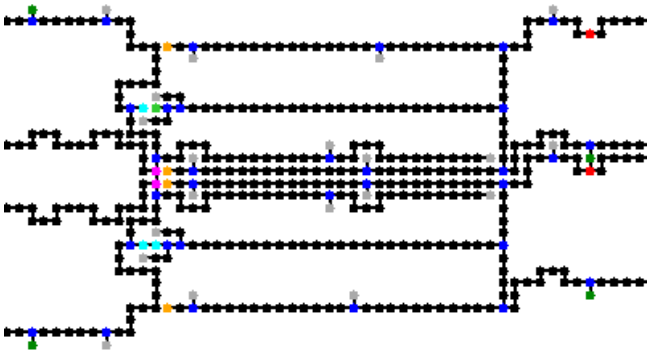


Figure 10: The central part of $\mathrm{P_{01}D_{01}}$, showing the way in which the polymer ($\mathrm{P_{01}}$) part on the right 'attacks' and weakens the dimer ($\mathrm{D_{01}}$) part on the left at the bond between the two magenta Perturbium Atoms. The two parts of $\mathrm{P_{01}D_{01}}$ are held together temporarily by the weak bond between the two cyan Anti-Loosium Atoms.

The dimer and polymer parts are weakly bound at the Anti-Loosium/Anti-Loosium bond that joins the Locator of the $\mathrm{M_{01}}$ part of $\mathrm{P_{01}}$ to the bottom right of $\mathrm{D_{01}}$. The memory mechanism relies on $\mathrm{D_{sp}}$ being split by $\mathrm{P_{sp}}$, and not by any other polymer. 'Incorrect' splitting is ruled out because the two Half Probes on $\mathrm{P_{sp}}$ must be an exact match for the two Half-Acceptors on $\mathrm{D_{01}}$.

The dimer weakening occurs because the two (orange) Metal Atoms at the end of the two Half-Probes on the polymer are close to the two (magenta) Perturbium atoms on the dimer. This weakens the bond between them, and the dimer splits. The top ($\mathrm{M_{01}}$) part of the dimer falls away because it has no other bond either with the polymer or with the other ($\mathrm{M_{01}}$) part of the dimer. The other part of the dimer also spits from the polymer shortly afterwards, because the Anti-Loosium/Anti-Loosium bond holding the two together is weak, and so can only be temporary.

To summarise: A $\mathrm{P_{sp}}$ polymer binds temporarily to a $\mathrm{D_{sp}}$ dimer, and as a result the dimer is weakened. Both parts of the dimer separate from the polymer, which is then free to split another dimer. A dimer can only be split by the 'correct' polymer because the Probe and Acceptor formed by the Half-Probes and Half-Acceptors of the monomers involved must have compatible shapes.

Results

Preliminary 'proof of concept' tests have been undertaken to investigate the workability of the memory bank described above. The tests used the SimSoup artificial chemistry simulator. Reactions take place in a well stirred Reactor. The rate constant k for Constructions is set to a constant value; those for Fissions are set to $k = Ae^{-E_f/RT}$, where E_f is the total energy of the bonds that have to be broken, T is temperature, and A and R are constants.

Results of two runs are presented. Both demonstrate memory; the first is typical of runs undertaken, the second illustrates an unusual 'ringing' phenomenon.

Run 1

The scenario for Run 1 is as follows:

- Starting at time 1000, a constant supply of 'food' is provided to a the Reactor. This consists of 400 Molecules of $\mathrm{M_{00}}$ every ten timesteps, plus 200 Molecules of each of $\mathrm{D_{01}}$, $\mathrm{D_{02}}$ and $\mathrm{D_{02}}$ every ten timesteps
- 'Seed' Molecules are added as follows: Five Molecules of $\mathrm{M_{01}}$ at time 10000, five Molecules of $\mathrm{M_{02}}$ at time 30000, five Molecules of $\mathrm{M_{03}}$ at time 50000
- At each timestep, every Molecule has a probability of 0.001 of being removed from the Reactor ('leakage')
- The size of Molecules was limited. This was necessary to enable the simulation to run within a reasonable time[8]

Figure 11 shows the numbers of the three polymers $\mathrm{P_{01}}$, $\mathrm{P_{02}}$ and $\mathrm{P_{03}}$ present in the Reactor at each timestep, along with the number of $\mathrm{M_{00}}$ Molecules[9].

[8]The operation of SimSoup is such that whenever a new Molecule Type enters the Reactor as a result of two Molecules joining, all the possible ways the new molecular ('board') structure can interact with existing molecular structures must be calculated. This is computationally intensive.

[9]$\mathrm{M_{00}}$ can be regarded as a polymer of length 1.

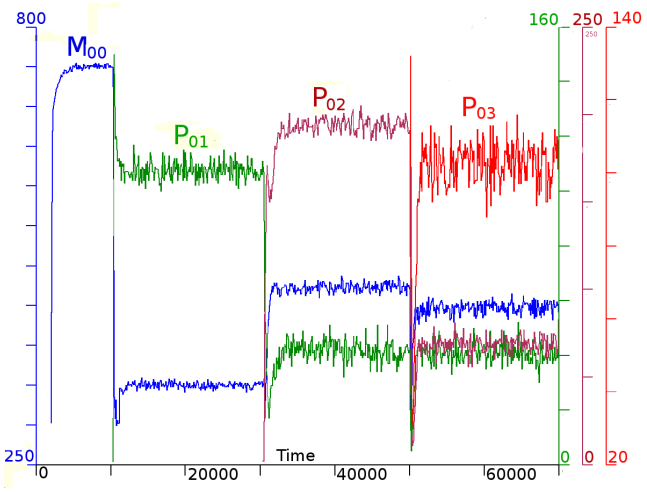


Figure 11: Plot showing the number of Molecules of M_{00}, P_{01}, P_{02} and P_{03} present in the Reactor during Run 1.

The addition of the 'seed' Molecules at times 10000, 30000 and 50000 in each case triggers a substantial change that persists over time. Prior to time 10000, there had been no Molecules of P_{01} present. Subsequent to the addition of the Molecules of M_{01} at that time, the number of P_{01} Molecules was stable at about 110 until time 30000.

Similar observations apply in regard to P_{02} and P_{03}. In each case, the seeding triggers a new persistent state in which the new Molecule Type is subsequently maintained.[10]

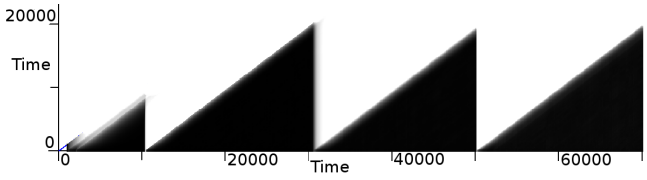


Figure 12: Manhattan Plot for Run 1. The black triangles indicate periods during which the Reactor composition (ie 'mix' of Molecule Types) varies little. The right hand edges of the triangles indicate sharp changes in composition.

Figure 12 is a 'Manhattan Plot' showing how the overall Reactor composition varied during the run. The construction of the Manhattan Plot has been described elsewhere (Gordon-Smith, 2007). The black triangles indicate periods during which there is little change in the composition (or 'mix') of Molecules in the Reactor.

The plot indicates that the pattern shown in Figure 11 in relation to a few key Molecule Types occurs more generally for the Reactor composition as a whole. There are periods of roughly constant composition, and sharp changes corresponding to the addition of the 'seed' Molecules.

The number of Molecule Types present in the Reactor (not shown) was high; at the end of the run it was almost 500.

[10]The numbers of Molecules of existing types has a step change each time a new state is entered. This is to be expected since the overall dynamics of Interactions in the Reactor are changed. However, this does not lead to the disappearance of an existing type.

Run 2

The scenario for Run 2 is similar to that for Run 1. There are differences in the timings at which Molecules are added.

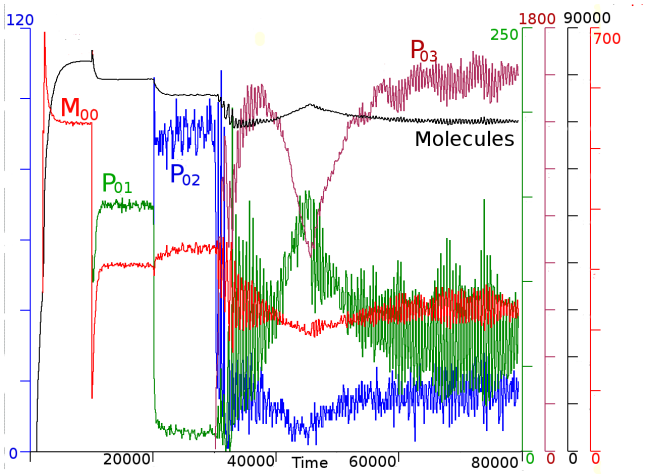


Figure 13: Time series plot for Run 2, showing 'ringing'.

Figure 13 is a time series plot for Run 2. 'Seed' Molecules are added at times 10000, 20000 and 30000. The system 'remembers' each seeding as for Run 1 in Figure 11. However, after the third seeding the system shows a variable oscillatory or 'ringing' behaviour before stabilising.

Discussion

Stability: The stability of the active state of a memory unit derives from the positive feedback mechanism that it incorporates. The design strategy for Molecule Types to support feedback is as follows. Firstly, identify each memory unit with a short polymer P_{sp} that can be produced from $P_{s,p-1}$ by the addition of an M_{sp} monomer, and which can catalytically split a closed dimer D_{sp} to produce more M_{sp} monomers. Then design the monomers to join only in ways that lead to production of the 'correct' polymers, and ensure that these polymers only split the 'correct' closed dimers.

Transition of a memory unit from the inactive to the active state can be triggered by addition of just a single monomer. A suppression mechanism could be added if necessary for stability, although this would add model complexity.

Moderate Complexity Of Monomers: The designed monomers are moderately complex, although far below the complexity of DNA and RNA and the molecules involved in their replication. There may be scope for simplification. It can also be envisaged that they could be products of some

systematic process that would result in co-ordination of the positions of the various projections and recesses.

Bias In Direction Of State Changes: Changes in state of a Memory Bank in an organism only take place in the direction of increasing p. However, this does not mean that evolutionary 'mistakes' cannot be reversed. If an organism is less fit as a result of a mutation then it will be less likely to persist in future generations. It may be possible to change the design of the molecules to remove the bias, but it does not rule out open-ended evolution at the ecosystem level.

Integration With Larger Network: Although the Memory Bank consists of a number of independent rows (or series), it can be envisaged to be integrated within a larger metabolic network that it influences.

Conclusions And Prospects

Conclusions

- An artificial chemical network and associated molecular structures designed to support up to 10^{10} persistent states has been shown[11]. It is reasonable to suppose that this would be sufficient for open-ended evolution to begin
- The monomers are of only moderate complexity, supporting the view that molecules with similar capabilities and properties (though no doubt very different structure) were present in the prebiotic world
- The operation of a small set of memory units has been simulated, completing the first part of the proof of concept
- Supplementary material for this paper is available at http://www.simsoup.info/Publications.html

Prospects

- It will be appropriate to make optimisations enabling a larger set of Memory Units to be tested
- The author would like to hear from anyone interested in translating the ideas described here to 'real' chemistry.

References

Cairns-Smith, G. (1982). *Genetic takeover*. Cambridge University Press.

Dittrich, P. and di Fenizio, P. S. (2007). Chemical organisation theory. *Bull. Math. Biol.*, 69(4):1199–1231.

Dyson, F. (1999). *Origins Of Life*. Cambridge University Press.

Fernando, C. and Rowe, J. (2007). Natural selection in chemical evolution. *J. Theor. Biol.*, 247:152–167.

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

Gordon-Smith, C. (2005). SimSoup: An artificial chemistry model for investigation of the evolution of metabolic networks. ECAL Workshop. http://www.simsoup.info/Publications.html.

Gordon-Smith, C. (2007). Evolution without smart molecules. ECAL Workshop. http://www.simsoup.info/Publications.html.

Gordon-Smith, C. (2009a). The origin of life: A network oriented view. In *Proceedings - Levels of Selection and Individuality in Evolution: Conceptual Issues and the Role of Artificial Life Models, September 2009*.

Gordon-Smith, C. (2009b). SimSoup: Artificial chemistry meets Pauling. In *Advances in Artificial Life: 10th European Conference, ECAL 2009, Proceedings*, Lecture Notes in Computer Science. Springer-Verlag.

Jain, S. and Krishna, S. (1998). Autocatalytic sets and the growth of complexity in an evolutionary model. *arXiv:adap-org/9809003v1*.

Kauffman, S. A. (1993). *The Origins Of Order*. Oxford University Press.

Krishna, S. (2003). *Formation And Destruction of Autocatalytic Sets in an Evolving Network Model*. PhD thesis, Indian Institute Of Science.

Oparin, A. (1957). *The Origin Of Life On The Earth*. Oliver And Boyd.

Pauling, L. (1960). *The Nature Of The Chemical Bond*. Cornell University Press.

Segré, D., Ben-Eli, D., Deamer, D. W., and Lancet, D. (2001a). The lipid world. *Origins Life Evol. B.*, 31:119–145.

Segré, D., Lancet, D., Kedem, O., and Pilpel, Y. (1998). Graded autocatalysis replication domain (GARD). *Origins Life Evol. B.*, 28:501–514.

Segré, D., Shenhav, B., Kafri, R., and Lancet, D. (2001b). The molecular roots of compositional inheritance. *Journal Of Theoretical Biology*, 213:481–491.

SimSoup (2011). Simsoup artificial chemistry simulator source code. http://code.google.com/p/simsoup/.

Wächtershäuser, G. (1990). Evolution of the first metabolic cycles. *Proc.Natl. Acad. Sci. USA*, 87:200–204.

Wächtershäuser, G. (1997). The origin of life and its methodological challenge. *J. Theor. Biol.*, 187.

Wächtershäuser, G. (2006). From volcanic origins of chemoautotrophic life to bacteria, archaea and eukarya. *Philos Trans R Soc London [Biol]*, 361:1787–1808.

[11] The s and p indices both vary from 0 to 9. This gives 10 series each consisting of polymers of length between 1 (food) and 10.