A Functional Self-Reproducing Cell in a Two-Dimensional Artificial Chemistry

Tim J. Hutton
Biomedical Informatics Unit, Eastman Dental Institute for Oral Health Care Sciences,
University College London, 256 Grays Inn Road, London WC1X 8LD, UK
T.Hutton@eastman.ucl.ac.uk

Abstract

We show how it is possible to make a self-reproducing cell in an artificial chemistry by surrounding a replicating molecule with a semi-permeable membrane. The molecule can carry an arbitrary amount of information, encoded in a material form as a sequence of bases, as in DNA. The cells produce enzymes through a decoding of their base sequence, and these enzymes trigger reactions essential to the cell's survival. Earlier work in a similar artificial chemistry showed that replicators free in solution could obtain no survival advantage from producing enzymes; here we show that when surrounded by a membrane the replicators *can* obtain an advantage. We show that the cells reliably reproduce over many generations under environmental pressure for resources. By creating cells in a material-based artificial chemistry we hope that the system might have the potential for open-ended, creative evolution.

Introduction

A deeper understanding of the requirements for creative, open-ended evolution could be obtained if we were able to recreate it for experimental purposes. This remains an important open problem in artificial life (Bedau et al., 2000) although information-carrying replicators have been created, often specifically for the purpose, in many different media, including cellular automata, machine-code, artificial chemistries and wet chemistry.

Open-ended evolution is defined as the unbounded appearance of adaptive activity, for which we have a test, suggested in (Bedau et al., 1998) and refined in (Channon, 2002). The term creativity, on the other hand, goes beyond this by requiring the appearance of innovative design solutions such as "mechanisms for sensing new aspects of [the] environment and for interacting with it in new ways ... and also for the very notion of individuality to change in radical ways (e.g. the evolution of multicellular organisms from unicellular ones)." (Taylor, 2001) Undergoing an evolutionary transition (Maynard-Smith and Szathmáry, 1995) such as the one mentioned would certainly be an example, no artificial system to date has demonstrated such evolutionary creativity.

It should be noted that systems with a predefined fitness function can give an insight into some aspects of evolutionary design creativity (Lenski et al., 2003) but are of limited use for understanding how to build systems that exhibit open-ended creativity.

Taylor (Taylor, 1999; Taylor, 2001) discusses open-ended evolutionary creativity in depth and argues that, for it to be possible, the replicators must a) be fully embedded in their arena of competition, b) have rich, unlimited interactions between each other and with their environment, c) initially replicate implicitly, rather than using some encoding of the replication process, and d) be constructed entirely of 'material' components, allowing the possibility of different encodings of information. These suggestions give a possible reason why creative, open-ended evolution has not yet been demonstrated in an artificial system: no system to date satisfies all four requirements, with the theoretical exception of (von Neumann, 1966). In this paper we give details of a system that extends (Hutton, 2002) and discuss whether it satisfies all of Taylor's requirements.

Background

As a computationally less-expensive alternative to molecular dynamics simulation, artificial chemistries (ACs) provide a way of modelling chemical processes abstractly. Recent spatially-explicit AC models include lipid micelles and membranes (Madina et al., 2003) and template-replicating molecules (Hutton, 2002; Smith et al., 2003).

In (Ono and Ikegami, 2002), self-reproducing cells were demonstrated to have the capacity to evolve through the selection of catalysts that affect the formation of the lipid membranes. However, the evolutionary capacity in this system was very limited since little information could be inherited.

In (Hutton, 2002), a set of reactions in an AC was shown that allowed molecules carrying any amount of information to replicate. These molecules qualify as units of evolution and unlimited hereditary replicators (Szathmáry and Maynard-Smith, 1997) but failed to demonstrate any interesting behaviour because they could do nothing but replicate - their phenotype was minimal.

In (Hutton, 2003a), a method for the replicating molecules

to produce reaction-specific *enzymes* was presented, allowing the molecules some way of controlling their chemical environment. However, it was found that in the absence of a surrounding membrane the enzymes could confer no evolutionary advantage to the molecule that produced them, since they would diffuse away. Here we extend this work by showing how it is possible to put each molecule inside a semi-permeable membrane, allowing the molecule to retain sole use of the enzymes it produces and thus potentially to derive from them some survival advantage to outweigh the extra replication time required to produce them.

To continue to replicate successfully, the molecule must cause the membrane to grow and divide, and ensure that a copy of the molecule is present in each daughter cell. We show how this can be achieved by connecting the molecule to the membrane temporarily, as is common in bacterial reproduction. Ideally the molecules themselves would engineer this kind of functionality, by producing the enzymes that trigger the required reactions but this is not yet done because of the large size of the resulting cell. We evaluate the potential of the cells to demonstrate the evolutionary growth of complexity (McMullin, 2000) and whether such a system is capable of evolutionary creativity.

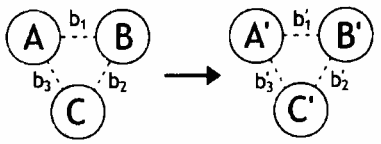
System Description

Our AC extends that of (Hutton, 2002), where 'atoms' have a fixed *type* $\in$ {a, b ... f} and a variable *state* $\in \{0,1,2 \ldots\}$. The atoms move around a finite world and a set of reactions (Table 1) determine what happens when they bump into each other; bonds between them may be formed (to make 'molecules') or broken and their states can be changed. Bonded atoms are prevented from moving apart.

Previously a lattice-based physics was used for speed, however this can lead to inflexibility because some of the atoms in large molecules will be unable to move. One solution is to allow bonded atoms to move further apart on the lattice but then it becomes difficult to make useful membranes since the contents of a bonded loop of atoms will be able to escape through the gaps.

A continuous-space physics (see eg. (Smith et al., 2003)) has better flexibility properties (at greater computational cost) and allows us to make membranes that prevent their contents from escaping, by using a simple spring force for volume exclusion. Semi-permeability is achieved by turning off the volume exclusion force for unbonded atoms in state 0 which can therefore pass through membranes (and other structures) freely. A more realistic model of lipid molecules might avoid the need for such measures at still greater computational cost.

Reactions 1-24 permit the molecule to make a copy of itself and cause the membrane to divide with one copy in each half. Figure 1 shows the various stages in the process, with the atoms in state 0 shown as dots and other atoms as filled circles (for clarity). The starting point is a loop of a8's, with a molecule such as e9-a1-b1-c1-d1-f1 inside, although importantly any sequence of a1's, b1's, c1's and d1's with an e9 at one end and an f1 at the other will replicate also. The molecule first attaches the 'e' end to the membrane, and the division ends with the molecules being released back into the daughter cells. Modern eukaryotic cells do not connect the strands of DNA to the membrane while duplicating but this is common in bacteria and appears to be a simpler way of achieving cell division.



	A	b_1	B	b_2	C	b_3	A'	b'_1	B'	b'_2	C'	b'_3
1	e9	✗	a8				10	✓	8			
2	e10	✗	e0	✗	a8	✓	4	✓	3	✓	8	✓
3	*x*3	✓	*x*4	✓	*y*1	✗	4	✓	4	✓	2	✓
4	*x*2	✗	*x*0				3	✓	5			
5	*x*3	✓	*x*5	✗	*y*4	✓	4	✓	3	✓	4	✗
6	f4	✓	f3	✗	a8	✗	7	✓	8	✓	10	✓
7	a10	✓	a8	✗	f8	✓	10	✓	11	✗	7	✗
8	*x*6	✓	*y*4				1	✓	7			
9	*x*7	✓	*y*7				6	✗	6			
10	e6	✓	a8				1	✓	8			
11	a8	✓	e1	✓	a12	✗	9	✗	1	✗	9	✓
12	a9	✓	a9	✓	a10	✗	10	✓	9	✗	14	✓
13	a14	✓	a10	✓	a8	✗	15	✗	14	✗	8	✓
14	a14	✓	e1	✗	a9	✓	8	✗	16	✗	8	✓
15	a15	✓	e1				8	✗	16			
16	e16	✗	f1				17	✗	17			
17	e17	✓	*x*1				17	✓	2			
18	e17	✓	*x*3	✓	*y*5	✗	17	✓	4	✓	3	✓
19	f17	✓	*x*4	✓	*y*3	✗	1	✓	6	✗	6	✓
20	*x*6	✓	e17	✓	*x*6	✗	18	✗	9	✓	1	✗
21	a10	✓	a11	✗	*x*1	✓	10	✓	12	✓	1	✓
22	*x*1	✓	*y*1	✗	a12	✓	1	✓	1	✓	13	✗
23	*x*1	✓	a13	✓	a10	✗	1	✗	11	✓	13	✓
24	*x*1	✓	*y*1	✓	a13	✓	1	✓	1	✓	10	✗
25	a8	✓	a7	✗	a8	✓	8	✓	8	✓	8	✗
26	*zi*	✓	*y*1				0	✗	*j*			
27	d*i*	✓	*y*1				*i*	✗	18			
28	d*i*	✗	*xp*	*u*	*yq*	✗	*i*	✗	*r*	*v*	*s*	✗

Table 1: The reactions used in our system. Reactions involve either two or three atoms, the column notation refers to the figure above. The types of the atoms are not given in the right-hand half of the table because they do not change. For example, the first reaction occurs when an e9 bumps into an a8 - the result is an e10 bonded to an a8. The symbols *x* and *y* are variables standing for any type a-f. Reactions 26-28 are for the production and application of enzymes and are explained in the text.

The membrane is pulled between the strands by reactions 21-24, in a sequence that passes a chain of atoms in state 1 along past a fixed point on the membrane. Such engineering solutions can also be seen in modern cells, for example the

microtubules on the mitotic spindle which act like conveyor belts or bargepoles to spatially organise the chromosomes and other components.

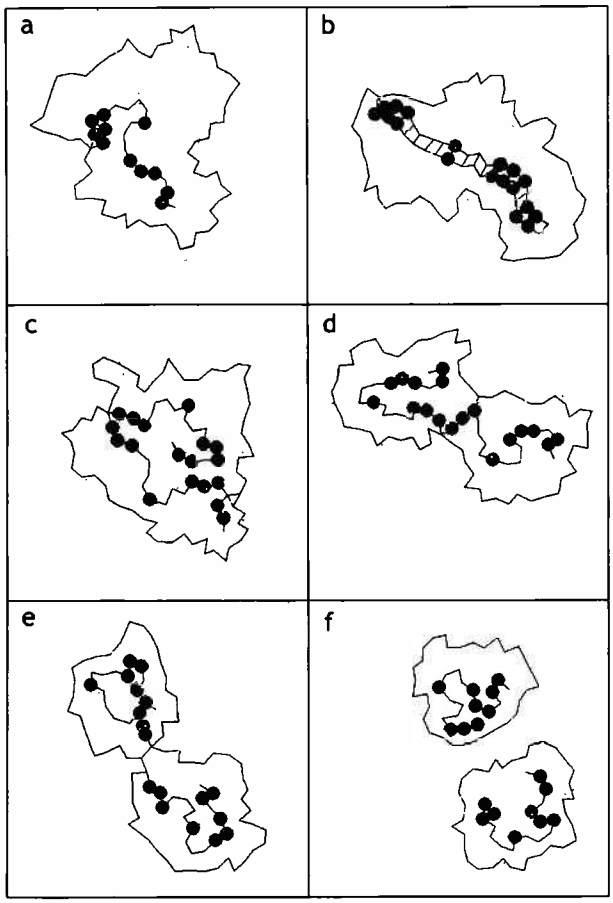


Figure 1: Six snapshots in the process of cell division in our artificial chemistry system. The information-carrying molecule first attaches itself to the membrane and duplicates itself (a, b). The membrane is pulled between the strands (c, d) before fusing at the top of the molecule (e) and dividing, releasing the molecules into the daughter cells (f) for the process to repeat.

There are several benefits to an information-carrying molecule of having a surrounding membrane, including a) the molecule is protected from any harmful external agents, and b) the molecule is able to control the chemical composition of its local environment. The second of these is achieved through reactions 26-28, which allow the production and application of enzymes.

The action of enzymes is implemented in reaction 28, which encodes any two-atom reaction involving states between `0` and `17` into a unique state i:

$$i = 2(2(6(6(18(18(18p+q)+r)+s) \\ +x)+y)+u)+v+18 \qquad (1)$$

where x and y are encoded as {`a`$=0, \ldots$ `f`$=5$} and u and v are 0 for unbonded and 1 for bonded. This formula is the most efficient way of encoding the values of p, q, r, s, x, y, u and v into a single number. For example, the enzyme `d5731` codes for `e0+a0`→`e2a3` and will trigger this reaction whenever the reactants bump into it.

Enzymes are produced through reactions 26 and 27. After each cycle of reproduction, the sequence of bases is turned from a non-reactive information sequence into an enzyme by a simple base-3 encoding given by reaction 26, where $i \geq 18$, z stands for any of the types `a`, `b` or `c`, and j is computed by:

$$j = 3(i-18) + \text{value}(z) + 18 \qquad (2)$$

with value(`a`) $= 0$, value(`b`) $= 1$, value(`c`) $= 2$.

This reaction converts a special copy of the base sequence, produced after cell division, for example `c18-a1-b1-d1`, into `a20-b1-d1`, then `b24-d1` and then `d37`. This base-3 'shift and add' procedure is a simple way of converting a string of `a`'s, `b`'s and `c`'s into a single number.

Atoms in nature cannot store such amounts of information, the 'atoms' in our system would perhaps be better termed 'proteins'. The idea is that enzymes can be encoded as small objects (rather than large, complex three-dimensional molecules) because their core functionality (triggering a specific reaction) is simple. This is just one of the simplifications that we are using to make a computationally-amenable AC system that supports life.

Figure 2 shows two cells in different stages of producing their enzymes. Each base in turn is released in state `0`, its information (its type) having been incorporated into the enzyme that is being produced. Reaction 27 allows more than one enzyme to be produced, for example by the sequence `cbadbbaccd` which would produce `d37` and `d134`.

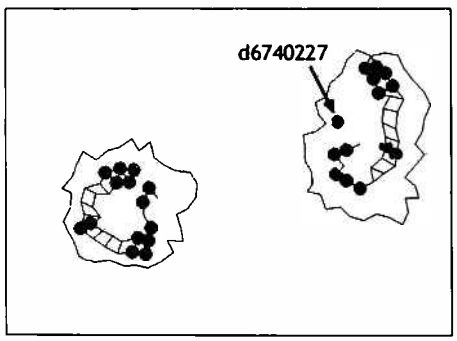


Figure 2: After cell division, the information-sequence is decoded into a reaction-specific enzyme. The sequence of bases is first duplicated (left) and then decoded into an enzyme (right). This particular enzyme (`d6740227`) causes the membrane to grow in length, helping the cell to continue to reproduce.

After division, the membranes of the daughter cells are approximately half the size of that in the original cell

(Fig. 1f), thus some mechanism for causing membranes to grow is necessary if reproduction is to continue. The cell achieves this by producing the necessary enzyme. In the cells shown the base sequence is `bbacaabaccbbbabd`, which is converted by repeated applications of reaction 26 into `d6740227` (Fig. 2), which triggers the reaction `a8+a0→a8a7`, adding extra atoms to the membrane. The atoms are incorporated into the membrane by reaction 25, making it bigger and allowing the cell to continue to reproduce. Cells without the correct base sequence to produce this enzyme or an equivalent one will run out of room inside the membrane and will cease to replicate.

Observations

Sometimes the cells depicted in Figs. 1 and 2 fail to divide evenly, leaving one daughter with insufficient room in the membrane to reproduce, this is one aspect of the design that could be improved, ideally by evolution. Also, there is no timing between the growth of the membrane and duplication of the information-sequence, this would also be desirable for the cell to reproduce more reliably. Even with these issues, the cells are capable of indefinite reproduction, as long as there are sufficient space and material resources.

To demonstrate this, we use the same experimental set-up as in (Hutton, 2002), where a fixed-size world is periodically 'flooded', with one half being cleared and replenished with raw material (unbonded atoms in state `0`). The reproducing entities have until the next flood to repopulate the empty half, and if they cannot do so quickly enough their numbers will drop to zero. Figure 3 shows the world after around 13 floods of 200,000 timesteps each, surviving cells are visible in amongst the broken components of cells cut in half by the flood.

Figure 3: A screenshot at approximately 2,600,000 timesteps, with floods every 200,000 timesteps. Some surviving cells are visible, amongst broken bits where the flood has cut cells in half.

In Fig. 3 it can be seen that the information molecule can survive contact with the outside chemical environment, it can also continue to replicate as long as it encounters strips of membrane. One side-effect of the cell division process is that these loose molecules can even produce complete cells again by a process of invagination; by wrapping a length of membrane around one copy of the molecule.

To investigate if this process is accounting for the continued survival of the cells, we rerun the experiment and introduce caustic agents into the external environment. These agents break bonds between any atoms they encounter, with the exception of the membrane atoms, making survival inside a membrane the only option. To test the evolutionary stability of the cells we introduce a low probability mutation reaction (p=0.0000001 per atom per timestep) that either adds or deletes a base, as R15 in (Hutton, 2003b). In this experiment we have used a flood that rotates between the four quarters of the world, allowing a smaller area to be simulated. Figure 4 shows a screenshot after 1,136,000 timesteps (flood every 50,000 timesteps), surviving cells are visible in amongst many lengths of severed membrane.

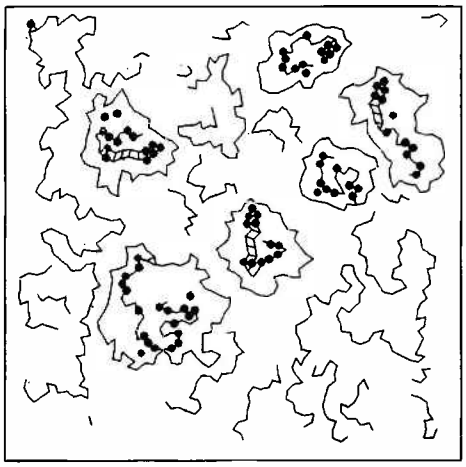


Figure 4: A screenshot after 1,136,000 timesteps, with a different quarter of the world being flooded every 50,000 timesteps. Six surviving cells are visible amongst the sections of membranes from cells cut in half by floods.

The need to keep the insides of cells separate from caustic agents in the external environment is seen in natural cells as well - if there were no barrier then the delicate machinery of reproduction would soon be perturbed. This is one of the main functions of a cell membrane.

By tracking the occurences of reaction 1 and reading off the string of bases connected to the 'e' atom we can observe which sequences are reproducing throughout the experiment. Figure 5 shows the reproduction rates for different sequences in a typical run, with the original sequence (`ebbacaabaccbbbabdf`) dominating, and mutants failing to reproduce for long. These results confirm that the

cells are capable of repeated reproduction, even without timing between the growth of the membrane and the replication of the information-carrying molecule.

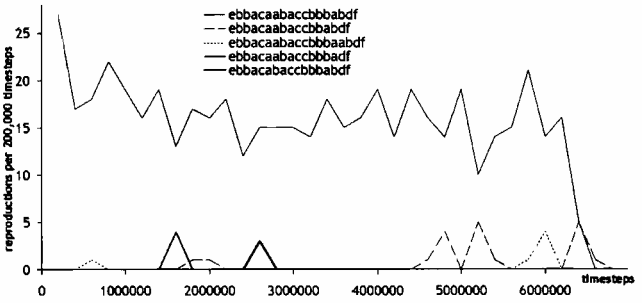


Figure 5: A plot of the reproduction rates for different cells in a typical run. The cell with the sequence necessary for membrane growth (`ebbacaabaccbbbabdf`) is the only one that continues to survive. Several mutations appeared in this particular run, with bases either missing or added. None survived for very long since they did not produce the correct enzyme (or an equivalent one) that helps the membrane to grow in size. Due to the small size of the area being simulated, the weight of mutations caused an extinction event at 6,600,000 timesteps.

The selection pressure against mutated sequences (Fig. 5) shows that carrying a significant amount of information can be beneficial in an AC system of reproducing entities. In this system the information is useful because the enzyme it produces is retained for the exclusive use of the cell. This was not the case in previous work when the molecules didn't have a surrounding membrane, instead the molecules soon evolved to shed their bases, even if the enzymes were required for reproduction (Hutton, 2003a). In the current system the cell shown is the smallest possible (since the enzyme that causes the membrane to grow is essential to reproduction), however it is hoped that the cells might evolve the production of *additional* enzymes to help them reproduce more reliably, or faster.

Conclusions

We have demonstrated an artificial chemistry (AC) containing a simple form of cell. An arbitrary sequence of bases encodes information for producing enzymes that can assist in the cell's reproduction. The cell is functional in the sense that it is capable of doing things other than just reproduce, through the action of its enzymes. We have shown that the cell is capable of robust, repeated reproduction in a changing chemical environment. The genetic and membrane subsystems are tied together, each playing a role in the production of the other. This form of cooperation has been discussed at length in designs for synthetic cells in wet chemistry (Rasmussen et al., 2003) and in abstractions of the life processes in natural cells (Ganti, 1997).

Taylor's requirements for a system to exhibit open-ended, creative evolution (Taylor, 2001) seem to be satisfied in this system, with one exception. The cells are information-carrying reproducers, and are fully embedded in their environment. Currently the cells reproduce implicitly (the reproduction rules are given) although since they are encoded as material components the potential exists for reproduction to proceed in a different manner, perhaps by using a set of enzymes. However, the cells do not currently interact in a very rich way since they have no way of sensing each other's presence. This is potentially a problem for evolvability, since without rich interactions there is no drive towards higher sophistication. The cells as they are could compete in a passive way, by making their reproduction cycle more robust or rapid by producing additional enzymes but a great number of enzymes would be required to make structures for defence or aggression. For comparison, in natural cells the number of enzymes required just for growth and division has been estimated at between 2,000 and 5,000 (Murray and Hunt, 1993). One possibility for allowing interactions between cells is through cell signalling - if some enzymes could occasionally move through the membrane then simple messages might be usefully passed. Cell signalling is an important prerequisite for cell cooperation and multicellularity.

Many of the reactions in Table 1 could be encoded as enzymes and produced by the cells themselves. This would reduce the number of required reactions and increase the size of the smallest viable cell. While this would make the first replicator less likely to appear spontaneously it might improve the evolvability of the system. The three-way reactions could be replaced by a set of two-way reactions, although this would require additional states, and in theory a system with only reactions 26, 27 and 28 (or even less) could support self-reproducing cells.

The earliest proto-cells on Earth are likely to have been formed through encounters between lipid aggregates such as vesicles, and autocatalytic polymers such as RNA. Understanding how these entities could have become a single unit, with the components helping each other's survival, is an important goal in understanding the origin of life (Ganti, 1997; Szostak et al., 2001). This study is not directly aimed at this question but draws inspiration from speculations about early forms of life, and may help eliminate hypotheses, as long as the abstractions used in the simulation are justifiable in the context.

The current system is implemented in two dimensions to minimize the computational costs and make visualisation easier but there might be advantages to moving to a three-dimensional physics, due to the greater number of topological possibilities. The reaction rules for template-replication could remain the same in 3D but a different approach to membrane division would be required. Solutions seen in nature that could be explored include contractile rings (tightening a belt around a balloon) and septation (growing a sheet

across the middle).

A more realistic simulation of lipids might reduce the complexity required to design a working AC cell, or rather transfer some of the design complexity into the environment. However, it seems that such an approach would necessarily incur a greater computational cost, and this is a major problem when trying to recreate an evolutionary system, since a fundamental tenet is that there must be many interacting entities.

The use of a semi-permeable membrane around an information-carrying replicator solves the problem encountered in (Hutton, 2003a) and discussed in (Szostak et al., 2001) and elsewhere, that a naked replicator can obtain no survival advantage from the production of enzymes. It is speculated that cells that additionally produce other enzymes might outcompete the ones presented here, and thus that the system might evolve to greater complexity. By following the principles of (Taylor, 2001) we hope that the system might eventually exhibit open-ended, creative evolution.

The source code for the implementation presented here is available at: `http://www.eastman.ucl.ac.uk/~thutton/Evolution/Squirm3/`

Acknowledgments

The author would like to thank Dan LaLiberte and Chrisantha Fernando for much interesting discussion and motivation. The anonymous reviewers were also very helpful in clarifying the important points of this paper.

References

Bedau, M., McCaskill, J., Packard, N., Rasmussen, S., Adami, C., Green, D., Ikegami, T., Kaneko, K., and Ray, T. (2000). Open problems in artificial life. *Artificial Life*, 6(4):363–376.

Bedau, M., Snyder, E., and Packard, N. (1998). A classification of long-term evolutionary dynamics. In Adami, C., Belew, R., Kitano, H., and Taylor, C., editors, *Proc. Artificial Life VI*, pages 228–237. MIT Press.

Channon, A. (2002). Improving and still passing the ALife test: Component-normalised activity statistics classify evolution in Geb as unbounded. In Standish, R., Bedau, M., and Abbass, H., editors, *Proc. Artificial Life VIII*, pages 173–181. MIT Press.

Ganti, T. (1997). Biogenesis itself. *J. Theor. Biol.*, 187(4):583–593.

Hutton, T. (2002). Evolvable self-replicating molecules in an artificial chemistry. *Artificial Life*, 8(4):341–356.

Hutton, T. (2003a). Information-replicating molecules with programmable enzymes. In *Proc. Sixth International Conference on Humans and Computers*, pages 170–175, University of Aizu, Aizu-Wakamatsu, Japan.

Hutton, T. (2003b). Simulating evolution's first steps. In Banzhaf, W., Christaller, T., Dittrich, P., Kim, J., and Ziegler, J., editors, *Proc. Seventh European Conference on Artificial Life*, pages 51–58, Dortmund, Germany.

Lenski, R., Ofria, C., Pennock, R., and Adami, C. (2003). The evolutionary origin of complex features. *Nature*, 423:139–144.

Madina, D., Ono, N., and Ikegami, T. (2003). Cellular evolution in a 3D lattice artificial chemistry. In Banzhaf, W., Christaller, T., Dittrich, P., Kim, J., and Ziegler, J., editors, *Proc. Seventh European Conference on Artificial Life*, pages 59–68, Dortmund, Germany.

Maynard-Smith, J. and Szathmáry, E. (1995). *The major transitions in evolution*. Morgan-Freeman.

McMullin, B. (2000). John von Neumann and the evolutionary growth of complexity: Looking backwards, looking forwards... *Artificial Life*, 6(4):347–361.

Murray, A. and Hunt, T. (1993). *The cell cycle: an introduction*. Oxford University Press.

Ono, N. and Ikegami, T. (2002). Selection of catalysts through cellular reproduction. In Standish, R., Bedau, M., and Abbass, H., editors, *Proc. Artificial Life VIII*, pages 57–64. MIT Press.

Rasmussen, S., Chen, L., Nilsson, M., and Abe, S. (2003). Bridging nonliving and living matter. *Artificial Life*, 9(3):269–216.

Smith, A., Turney, P., and Ewaschuk, R. (2003). Self-replicating machines in continuous space with virtual physics. *Artificial Life*, 9(1):21–40.

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.

Szostak, J., Bartel, D., and Luisi, P. (2001). Synthesizing life. *Nature*, 409:387–390.

Taylor, T. (1999). *From artificial evolution to artificial life*. PhD thesis, Division of Informatics, University of Edinburgh.

Taylor, T. (2001). Creativity in evolution: Individuals, interactions and environment. In Bentley, P. and Corne, D., editors, *Creative Evolutionary Systems*, chapter 1. Morgan Kaufman.

von Neumann, J. (1966). *Theory of self-reproducing automata*. Urbana: University of Illinois Press. Edited and completed by Arthur Burks.