

Stochastic innovation as a mechanism by which catalysts might self-assemble into chemical reaction networks

Justin A. Bradford[†] and Ken A. Dill^{*§}

[†]Graduate Group in Biophysics and ^{*}Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143

Communicated by Christian de Duve, Christian de Duve Institute of Cellular Pathology, Brussels, Belgium, April 16, 2007 (received for review June 29, 2006)

We develop a computer model for how two different chemical catalysts in solution, A and B, could be driven to form AB complexes, based on the concentration gradients of a substrate or product that they share in common. If A's product is B's substrate, B will be attracted to A, mediated by a common resource that is not otherwise plentiful in the environment. By this simple physicochemical mechanism, chemical reactions could spontaneously associate to become chained together in solution. According to the model, such catalyst self-association processes may resemble other processes of "stochastic innovation," such as Darwinian evolution in biology, that involve a search among options, a selection among those options, and then a lock-in of that selection. Like Darwinian processes, this simple chemical process exhibits cooperation, competition, innovation, and a preference for consistency. This model may be useful for understanding organizational processes in prebiotic chemistry and for developing new kinds of self-organization in chemically reacting systems.

chemical evolution | self-organization | abiogenesis | catalytic chains

There are several examples of what might be called "stochastic innovation," whereby a biological, physical, or sociological system: (i) searches among viable options, then (ii) selects one or more of those options that is "best" by some metric, then (iii) locks in that selection for the future. In biology, the best-known example is Darwinian evolution (1), where variation is the term that describes the search step, and natural selection is the term that describes steps (i) and (ii). Stochastic innovation occurs in learning and memory and neural development, where the correlated firings of neurons can lead to changes in synaptic strength (2–5), and to the development of the vascular system, where new blood vessels grow toward oxygen-deprived cells (6, 7).

Stochastic innovation appears in other arenas, too. Human beings, businesses, and social organizations evolve through decision-making: they search among the options available to them, make self-serving choices, then remember and act on those decisions in the future. Social insects, like ants, search randomly for food, then lock in the discovery with chemical trails for the rest of the colony (8, 9). Computer models of artificial life and artificial societies show that stochastic innovation can meet apparent goals that were not programmed into them at earlier times (10, 11). The power of stochastic innovation is that it can lead to complex behaviors or organizational structures that are responsive to changes in the environment, even though such processes are unguided, unplanned, and stochastic.

Our interest here is in whether stochastic innovation might also be achievable in chemistry and biochemistry. Can chemical and biochemical reactions be chained together in complex and innovative ways, driven only by simple physicochemical search and selection processes? If so, it may be useful, not only as a tool in chemistry and biochemistry, but also for giving insights into the processes of chemical organization that may have occurred during prebiotic evolution.

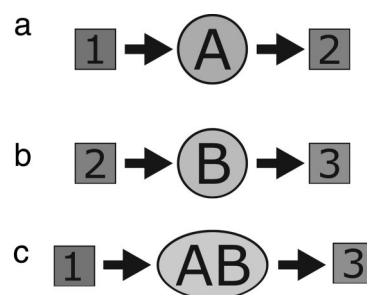


Fig. 1. Agents (lettered circles) and resources (numbered squares). (a) Agent of type A converts substrate 1 to product 2. (b) Agent of type B converts substrate 2 to product 3. (c) When agents A and B are complexed together, two reactions are chained together, converting substrate 1s to product 3s.

Here, we propose a simple model. Our goal is not to explain some existing body of data, because we know of none that pertains. Rather, our goal here is to propose a type of organizing principle that has not been explored before, as far as we know, but that is based on well established physicochemical principles and that can be tested by experiments. Our initial motivation for this work was to understand some puzzles of prebiotic chemistry, where, it could be argued, the field is just as limited by a lack of specific testable models at the moment as it is by a lack of experiments.

Model of Agents and Resources. We focus here on catalysts, such as enzymes or simple surfaces. We call a catalyst an "agent." An agent converts a substrate to a product; we label agents alphabetically (see Fig. 1). We call a substrate or product a "resource"; we label resources numerically. We assume that agents are Michaelis–Menten catalysts; i.e., they bind to their substrate before converting the substrate to a product. In our model, resources may be supplied by the external environment. Such environmental resources may vary with time in a random or controlled way by external forces, but, for simplicity, we assume they are uniformly distributed in space.

Fig. 1 shows an example. Agent A converts a substrate 1 to a product 2. Agent B converts substrate 2 to product 3. Key components of our model are the common resources, which are substrates or products that serve in common among different types of agents. For example, in Fig. 1, resource 2 is a common

Author contributions: J.A.B. and K.A.D. designed research; J.A.B. performed research; J.A.B. analyzed data; and J.A.B. and K.A.D. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: GF, genetics-first; MF, metabolism-first.

[§]To whom correspondence should be addressed. E-mail: dill@maxwell.ucsf.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0703522104/DC1.

© 2007 by The National Academy of Sciences of the USA

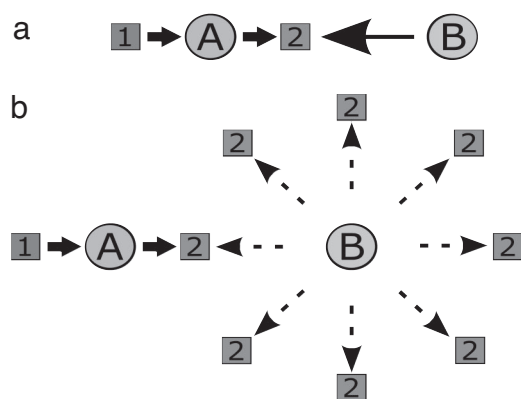


Fig. 2. Attraction and shielding. (a) Attraction. Bs are attracted to 2s, which are produced by As, hence Bs are attracted to As. (b) Shielding. When 2s are plentiful in the environment, Bs are attracted to them in all directions, hence have no special net tendency to associate with As.

resource because it is both a product of A and a reactant for B. Figs. 1 and 3 also show that if agents A and B come together by some process, then the AB complex is a “machine” that converts 1s to 3s, mediated by the intermediary resource 2s.

Agent B (Fig. 1) may take up a substrate molecule 2 from either of two sources: either the 2 was produced as the output from a nearby A agent, or the 2 was supplied externally from the environment, if 2s are available from external sources. Because we assume that Bs are Michaelis–Menten catalysts, a B will bind to its substrate, a 2 in this case. Bs will concentrate around 2s simply because Bs flow down their chemical potential gradients, in the same way that solutes in chromatographic mobile phases will seek out and bind to stationary-phase surfaces for which they have affinity.

Principles of Attraction and Shielding. There are two possibilities for each B agent (Fig. 2): attraction or shielding.

1. Attraction. As attract Bs through the following indirect mechanism: As produce 2s; 2s are localized near the As; those 2s attract Bs, concentrating the Bs around the As, thus leading to more AB complexation than would have occurred without the intermediary 2s (see Fig. 2a). This enhancement happens when: (i) A agents are present, (ii) 1s (the substrates for As) are plentiful, and (iii) 2s are depleted in the environment (i.e., available only at small or zero concentrations). AB complexation introduces into the system an “innovation,” i.e., an ability to produce 3s from 1s, an ability that does not simply and directly result from the presence of A or B alone (see Fig. 1c). Significant chaining together of agents is an emergent property of our system. In short, AB complexes are driven to form through mutual indirect attraction, mediated by 2s, the common resource, but this attraction occurs only when 2s are depleted from the environment.
2. Shielding. In contrast, when environmental 2s are plentiful and available in all directions, Bs will not selectively migrate toward As (see Fig. 2b). We call this “shielding.” (A more biological example of shielding is chemotaxis. A bacterium will swim toward a point source of food, except if food is uniformly distributed everywhere in space; then the bacterium would not migrate preferentially toward any one single point source. Chemotaxis, of course, is a complex process, but it illustrates how a favored direction of motion can result from simple physicochemical forces that change when environmental resources vary.) In the present model, attraction and shielding are simple consequences of concentration gradients.

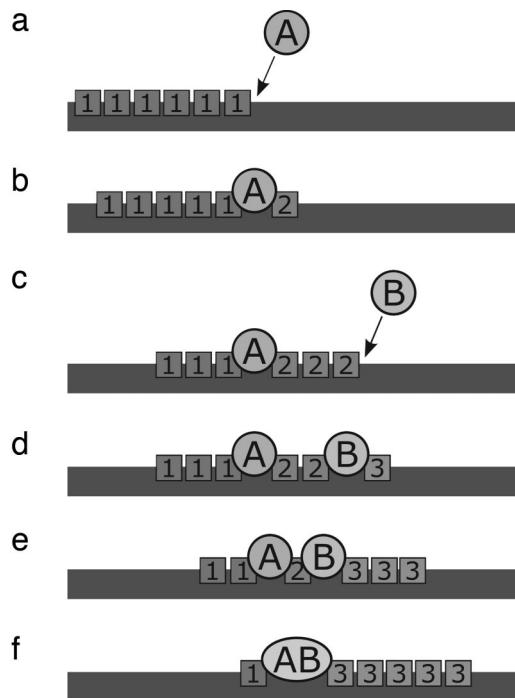


Fig. 3. Attraction on the lattice. (a) Agent A leaves the bulk and binds to the surface lattice in a region where resource 1s are concentrated. (b) Agent A converts 1s to 2s. (c) Agent B leaves the bulk and binds to the surface lattice where 2s are concentrated, which tends to be near the As that produced them. (d) Agent B produces 3s. (e and f) Agent B associates with A (e), forming a complex, which is now (f) a machine that converts 1s to 3s.

Details of the Model. Here is our model for how stochastic innovation might arise in a system of chemical catalysts. We assume that catalyst agents can adsorb to a surface. The prebiotic origin of life may have involved surfaces, such as minerals or clays, on which reactions took place (12–14).

Our model involves two compartments. First, there is a surface lattice where all of the reactions take place. Second, that surface is in direct contact with a bulk solution just above it, which serves as a chemical potential “bath,” a source of agents and resources for the surface. The surface simply provides a mechanism for trapping the 2s and Bs, slowing their escape from the As, providing the basis for the AB complexation enhancement. There are N_A molecules of each agent type in the simulation, and $E_N(E_1, E_2, \dots)$ represent the concentrations of that resource in the bulk solution. Fig. 3 shows the steps: (i) 1s attract As from the bulk onto the surface lattice, (ii) As produce 2s, (iii) which then attract Bs, (iv) leading to a machine in which As and Bs are clustered on the surface to produce 3s from 1s. Agents and resources diffuse rapidly throughout the bath, so they can be regarded as being in equilibrium within the bath over the time scale of the processes that happen on the surface. The lattice has $L_x \times L_y$ lattice sites; it just serves to coarse-grain the spatial localization. Each lattice site is large enough to be occupied by multiple molecules of different types at the same time.

We believe this is the minimal model that captures how catalysts might self-associate in solution. During any given time interval in the simulation, any or all of the following processes may occur (see Fig. 4).

1. A resource molecule or an agent molecule may drop down from the bulk and associate with the surface lattice, with rate coefficient k_{on} . To keep the model simple, all four types of molecule bind the lattice with the same rate constant.

many plausible models for the later stages, based on hypercycles (21, 22), and/or based on the many powerful RNA and protein evolution experiments that have been performed (23–25). The challenge in accepting GF as a first step, however, is that it is complex and requires the joint emergence of catalysis, compartmentation, and heritability, all at the same time (26). de Duve has noted (16) that GF “is accepted much less for its likelihood than for the lack of an alternative.”

In the MF model (14, 16, 27), chemical reactions become chained together evolutionarily before the appearance of genetic machinery. Although the requirements for MF are, in theory, more elementary than for GF, a key question about MF is how catalysts might become organized on their own, in the absence of a genetic system (28, 29).

We believe the present model of stochastic innovation based on attraction and shielding among chemical catalysts provides a plausible mechanism by which simple metabolic chains and cycles of reactions could have come together, perhaps at least long enough for a genetic system to then emerge. Of course, an important virtue of ultimately having a genetic system is that it provides much longer term “memory” for the “lock-in” step (step 3 in the Introduction) than does nongenetic propagation, where memory is merely provided by a ratio of off-rates to resource fluctuation times.

A key distinction between stochastic innovation, explored here, and design-based innovation, in which a complex system is engineered and constructed by a designer, is that stochastic innovation involves no implicit “goals” and no guidance toward a particular purpose. The Darwinian paradigm shows how increasing complexity and order can arise from processes that do not involve guidance through intelligence or design. Darwinian evolution is a process of elimination (“evolution-away-from”), rather than a process of design (“evolution-toward”). Stochastic innovation achieves evolution-away-from by search, selection, and lock-in. The present model has the three features of stochastic innovation.

1. Search. The B agents diffuse randomly through space and find 2s, based on a mutual binding affinity.
2. Selection. The B agents associate with A agents, if the common resource 2s are not present in the environment, leading to AB complex formation.
3. Lock-In. If the off-rate for dissociation of AB complexes is slower than the average frequency of resource depletion disasters of the common resource in the environment, then the complexes will be stable beyond the time scale of a single resource disaster.

Here is a possible experimental test of our model. Two enzymes, A and B, would be selected, based on having a common resource (e.g., 2). A’s product is known to be B’s substrate. The As would be covalently linked to a surface, such as a chromatography stationary phase. A pair of fluorescent probes would be attached: a donor probe is attached to each A molecule and an acceptor probe to each B molecule. This would allow for monitoring the concentration of AB complexes, through fluorescence quenching, for example. Two experiments would be performed. In one experiment, B molecules in a mobile phase would flow across the stationary phase in the absence of added 2s, but in the presence of 1s, the substrate for As, and then the amount of AB complex formation, would be measured. The other experiment would be identical, except that a high concentration of 2s would also be present in the mobile phase. The present model suggests that, if the relative on-rates, off-rates, and diffusion coefficients are roughly as given in **SI Table 1**, AB complexes should be more concentrated in the first experiment.

Related Modeling Efforts. The present model differs from others that have been used to explore chemical self-organization. Some agent-based models (10, 11, 30) involve computer-based rules rather than the laws of physical chemistry that are of interest here. Mass-action models of early evolution have been developed (31, 32), but they also are not focused on the microscopic chemical mechanisms. Models of hypercycles (21, 22) pertain to molecular systems that already have a genetic system, whereas our interest here is in molecular systems having no genetic system. Recently, there have been interesting studies of complexity and fragility in biological systems (33, 34), but those treatments presume some preexisting mechanism of stochastic innovation.

Because the work of Turing in the early 1950s (35), much mathematical modeling in chemistry, chemical engineering, and biology has explored pattern formation (36). Pattern formation is often modeled by using lattice models such as the present one, or nonequilibrium coupled spatiotemporal differential equations, sometimes with added stochastic noise terms. We believe that our model, too, could be readily cast in the form of such a coupled set of “reaction–diffusion” equations.

However, the present work differs from those treatments in certain respects. First, ours addresses not just how molecules can move and organize in space, but how chemical reactions can move and organize in space and in so doing become chained together into more complex reactions. Second, pattern formation usually involves some particular nonlinear component term in the model’s mathematics, often arising from saturable or cooperative binding, for example (37). Our model, too, involves saturable binding, but the most important nonlinear component of our model is what might be called “reflexive catalysis.” That is, at the same time that our agents catalyze reactions among the resources, our common resources also perform a sort of catalysis in the process of agent–agent association. In addition, our resources can be regarded as “ephemeral catalysts,” catalyzing complex formation only under certain nonequilibrium conditions and losing their catalytic power at equilibrium or when resources are plentiful.

Conclusion

A well known process in chemistry is the binding and association of molecules, driven by thermodynamic forces. Here, we consider whether catalyst molecules might be driven to associate with each other, through typical binding forces, but based on their molecular functions. Functional driving forces are well-known in biology, through the principles of evolution, but are not yet much studied in chemistry. We propose a model for how different Michaelis–Menten enzymes or catalysts might tend to associate, driven by the production or depletion of common resources. The agents do not associate if the common resource is plentiful. We call this the shielding principle. In this way, agents organize adaptively, and complexity can form from simpler systems. In our model, “function dictates structure,” a reversal of the paradigm in which “structure dictates function.”

Our model of stochastic innovation for chemical catalysts has three components: (i) a method for sampling a space of possibilities (Bs can search for As or not, depending on the presence of 2s), (ii) selection of a “favorable” outcome (Bs become more productive of 3s when their substrate 2s are depleted, if they associate with As), and (iii) a form of memory that locks in this outcome (the off-rates of Bs from As is slower than the time constant for fluctuations of 2s in the environment. Our process resembles other stochastic innovation processes, such as Darwinian evolution, in that agents compete, cooperate, value consistency, and innovate. The model is experimentally testable. It is a possible model for how metabolic reactions might have become chained together, at least primitively, before the for-

mation of a genetic system. And, it might lead to ways to self-assemble chemical reactions in solution.

Materials and Methods

Fig. 4 shows that B agents can associate with A agents through two routes: (i) slow random migration on the surface, after Bs bind to the lattice, or (ii) rapid diffusion in the bulk, then binding to the lattice. The latter is more targeted, because Bs bind preferentially (through parameter k_{coop}) where 2s are located, potentially near the As. Hence, the intrinsic rate of AB complexation depends on k_{on} , k_{off} , and D_{lattice} .

Intermediary 2s can affect AB complexation through two mechanisms. First, AB complexation is promoted if the number of lattice-bound 1s is high and if As are productive (k_{1-2} is large). The number of 1s on the surface depends on 1s in the bulk (E_1) and the adsorption equilibrium constant, $k_{\text{on}}/k_{\text{off}}$. The number of surface As is then determined by k_{coop} and dependent on the concentration of lattice-bound 1s. These bound molecules produce bound 2s at the rate k_{1-2} . However, if catalytic rates are much slower than k_{off} , resources disappear from the surface

before they can become concentrated enough to attract agents. Second, AB complexation is slowed by the environmental supply of 2s (E_2), which adsorb to the surface randomly and serve as decoys that attract Bs to nonproductive sites (where there are no As).

In order that 2s remain near As long enough to mediate AB complexation, lateral diffusion on the surface is set to be relatively slow, so dissociation of an AB complex on the surface is typically followed by reassociation. In contrast, dissociation of an AB complex in the bulk solution seldom leads to reassociation, because diffusion in the bulk is set to be fast. Therefore, most of the permanent dissociation of AB complexes occurs in the bulk solution. Also, the catalytic rates used in our simulations are large or comparable to the lateral diffusion rates.

We thank Patsy Babbitt, Don Hilvert, Tack Kuntz, Chris Voigt, John Chodera, and Yigal Nochomovitz for helpful discussions. This work was supported by a National Science Foundation predoctoral fellowship (to J.A.B.) and National Institutes of Health Grant GM34993 and the Sandler Foundation (to K.A.D.).

1. Mayr E (2001) *What Evolution Is* (Basic Books, New York).
2. Hebb DO (1949) *The Organization of Behavior* (Wiley, New York).
3. Kelso SR, Ganong AH, Brown TH (1986) *Proc Natl Acad Sci USA* 83:5326–5330.
4. Goda Y, Davis GW (2003) *Neuron* 40:243–264.
5. Alberts B, Johnson A, Lewis J, Raff, M, Roberts K, Walter P (2002) *Mol Biol Cell* (Garland Science, New York) p 1238.
6. Ferrara N, Gerber HP, LeCouter J (2003) *Nat Med* 9:669–676.
7. Kirschner MW, Gerhardt JC (2005) *The Plausibility of Life* (Yale Press, New Haven, CT) pp 143–176.
8. Goss S, Aron S, Deneubourg JL, Pasteels JM (1989) *Naturwissenschaften* 76:579–581.
9. Detrain C, Natan C, Deneubourg JL (2001) *Naturwissenschaften* 88:171–174.
10. Lenski RE, Ofria C, Pennock RT, Adami C (2003) *Nature* 423:139–144.
11. Epstein JM, Axtell R (1996) *Growing Artificial Societies: Social Science from the Bottom Up* (MIT Press, Cambridge, MA).
12. Hanczyc MM, Fujikawa SM, Szostak JW (2003) *Science* 302:618–622.
13. Segre D, Ben-Eli D, Deamer DW, Lancet D (2001) *Orig Life Evol Biosph* 31:119–145.
14. Wächtershäuser G (1988) *Microbiol Rev* 52:452–484.
15. Behe M (1996) *Darwin's Black Box* (Free Press, New York).
16. de Duve C (1987) *Proc Natl Acad Sci USA* 84:8253–8256.
17. Smith E, Morowitz HJ (2004) *Proc Natl Acad Sci USA* 101:13168–13173.
18. Orgel LE (1998) *Trends Biochem Sci* 23:491–495.
19. Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR (1982) *Cell* 31:147–157.
20. Joyce GF, Orgel LE (1999) *The RNA World*, eds Gesteland RF, Cech TR, Atkins JF (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) pp 49–77.
21. Eigen M, Schuster P (1982) *J Mol Evol* 19:47–61.
22. Eigen M (1971) *Naturwissenschaften* 58:465–523.
23. Bartel DP, Szostak JW (1993) *Science* 261:1411–1418.
24. Wright MC, Joyce GF (1997) *Science* 276:614–617.
25. Keefe AD, Szostak JW (2001) *Nature* 410:715–718.
26. Smith E, Morowitz HJ (2004) *Proc Natl Acad Sci USA* 101:13168–13173.
27. Kauffman SA (1986) *J Theor Biol* 119:1–24.
28. de Duve C (1991) *Blueprint for a Cell* (Neil Patterson, Burlington, NC) pp 123–170.
29. Orgel LE (2000) *Proc Natl Acad Sci USA* 97:12503–12507.
30. Troisi A, Wong V, Ratner MA (2005) *Proc Natl Acad Sci USA* 102:255–260.
31. Jain S, Krishna S (2001) *Proc Natl Acad Sci USA* 98:543–547.
32. Pfeiffer T, Soyer OS, Bonhoeffer S (2005) *PLoS Biol* 3:1269–1275.
33. Csete ME, Doyle JC (2002) *Science* 295:1664–1669.
34. Carlson JM, Doyle J (2002) *Proc Natl Acad Sci USA* 99:2538–2545.
35. Turing AM (1952) *Trans Roy Soc London* 237:37–72.
36. Murray JD (2002) *Mathematical Biology* (Springer, New York).
37. Strogatz SH (1994) *Nonlinear Dynamics and Chaos* (Addison-Wesley, Reading, MA).