Origins of Order in Dynamical Models

BRUCE H. WEBER
Department of Chemistry and Biochemistry
California State University
Fullerton
Fullerton, CA 92634–9480
U.S.A.


Understanding the emergence and evolution of biological order has been a fundamental goal of evolutionary theory ever since (even before) Darwin. Darwin sought to explain adaptation by the action of natural selection, out of which process order would emerge (Darwin 1859). David Depew and I have argued that Darwin was able to accomplish this, in part, through appeal to a metaphorical extension of models based upon Newtonian dynamics, and that later, through the work of Haldane, Fisher, Wright and Chetverikov, Darwinism was able to reformulate the concept of natural selection by appeals to dynamical models that were extensions of statistical mechanics and thermodynamics (Depew and Weber 1989; Depew and Weber 1995; Weber and Depew 1996). Currently there is debate as to whether natural selection as construed by the Modern Evolutionary Synthesis is sufficiently robust to account for large-scale biological order in addition to local adaptation. Defenders of the Synthesis argue that natural selection is still adequate and indeed the only alternative to regressive ideas of creationism or crypto-creationism (Ayala 1985; Williams 1992). Alternatively, modern-day proponents of an alternative research tradition to Darwinism, developmentalism, which reaches back to Geoffroy St. Hillaire, have argued that a principle other than selection is needed to complement or replace selection as the source of large-scale biological order (Saunders and Ho 1984; Goodwin 1989, 1994; Salthe 1993). Such modern-day developmentalists are making use of the current rapid development of methods of complex systems dynamics to argue that self-organization, rather than selection, is the dominant factor producing the order of biological systems (Brooks and Wiley 1988; Goodwin 1994; Salthe 1993).
There exists an alternative (among a number of such possibilities, see Weber and Depew 1996), in which both selection and self-organization act in concert to bring forth biological order. This position is explored by Stuart Kauffman. Kauffman, who was influenced both by John Maynard Smith and Brian Goodwin (representative figures in the opposed traditions of Darwinism and developmentalism), is one of an increasing number of theorists who seeks to develop a possible new evolutionary synthesis from within an extended Darwinian tradition. While it is too early to discern how successful such efforts will be or what the contours of such a synthesis would look like, we can make a preliminary evaluation of this program, as envisioned by Kauffman, and contemplate some of its implications.

Before addressing substance, some comments on Kauffman’s style are appropriate, especially in light of the criticisms of other reviewers (Dover 1993; Alberch 1994). This is a difficult book that has a complex, self-referential structure and that is not particularly “reader friendly.” Further editing might have helped. But a drastic reduction of the text by half (as has been suggested), while it might make the book more accessible, would also make it less interesting and useful. Like the subject matter itself, Kauffman’s book is best approached initially in a nonlinear manner, letting oneself be drawn into the parts that attract particular interest. On a second reading the structure and logic of the whole become more apparent. My experience is that graduate students and bright undergraduates cannot only understand this book but can be inspired by it to explore the “cyberspace” of dynamical models on their own. This book’s riches amply reward the effort of reading it.

Kauffman states that his goal is to broaden the context of Darwinism by incorporating self-organizational principles to evolutionary theory to reflect the fact that, while natural selection is always at work, it does not have to generate all the order of biological systems by itself. Implicit in this claim is the goal of achieving a synthesis of the Darwinian and developmentalist research traditions. Kauffman writes, “It is not that Darwin is wrong, but that he got hold of only part of the truth” (Kauffman 1993, xiii). Kauffman hopes through the use of the dynamical models of complex systems to forge a union of two streams of insight and so transform our understanding. In order “… to combine the themes of self-organization and selection, we must expand evolutionary theory so that it stands on a broader foundation and then raise a new edifice. … We must understand how such self-ordered properties permit, enable and limit the efficacy of natural selection” (Kauffman 1993, xiv, original emphasis).

It is true that Kauffman’s earlier work tended to emphasize the ways in which self-organization in complex networks limits the action of selection or provides the structures upon which selection could act (Kauffman
1969, 1974, 1985). At times these themes are revisited in *The Origin of Order*: “In sufficiently complex systems selection cannot avoid the order exhibited by members of the ensemble. Therefore, such order is present not because of selection but despite it” (Kauffman 1993, 16, original emphasis).

It is not germane that conservative Darwinians will doubtless view Kauffman’s work as non-Darwinian, for the main thrust of the book is to see self-organization and selection as interacting productively: “Thus the natural marriage of self-organization and selection first discovers the powerful order inherent in complex systems in the near-liquid regimes and then appeals to selection to achieve and sustain membership in this ensemble” (Kauffman 1993, 465, emphasis added). Kauffman argues for what he calls an “unrepentant holism” and a synthetic rather than an analytic biology, a biology that flows logically from his emphasis on complex systems properties (Kauffman 1993, 367).

Kauffman uses several types of models, the roots of which lie in the study of Boolean networks, to explore the dynamics of complex systems. Regardless of the model type, it needs to be emphasized that the models are about relationships abstractly considered. These models are in themselves neither biological, nor restrictively physical, but essentially mathematical. Nor, in and of themselves, do they imply any particular mechanism or apply to only a particular hierarchical level (contra Ulanowicz 1995, see also Rosen 1991). Indeed Kauffman’s models have been criticized for having just these properties and hence not being necessarily relevant to biological systems (Burian and Richardson 1991; Dover 1993). Kauffman’s NK model allows for a very general description of any system consisting of N components with K interactions between the components and in which there can be any number of states for each N. N could be the number of amino acids in a polypeptide chain, the number of polypeptides in a catalytic ensemble, the number of genes in a population, or a good deal else. These models are concretized by being mapped onto a fitness landscape such that a given protein sequence or gene in a population can be assigned a fitness value. For the purposes of modeling, fitness does not have to be defined, just the differences in relative fitness. With use of modern, high-speed computers, it is possible to simulate the dynamics of such systems as a response to both variation of parameters and over time. The reference to fitness landscapes provides much of the glue that bonds Kauffman’s work to the Darwinian tradition (specifically the Wright-Dobzhansky orientation).

What is provocative about such models is that within certain parameter ranges, they exhibit intrinsic and emergent ordering properties. We may distinguish three general cases: (1) When N is large and K is also large (up to K = N–1) the fitness landscape is rugged, with peaks about equal to
the mean fitness of the ensemble. (2) When \( K = 0 \), there is a smooth fitness landscape with only one peak of fitness that is much higher than the ensemble mean fitness (call this “Fisher world”). In the latter case the landscape is said to be \textit{correlated} in that a change in one component will make only a small change in fitness and will not affect the fitness of other components. In the former case, on the other hand, the landscape is said to be \textit{uncorrelated} in that a change in one component can result in a large change of fitness that can affect the fitness of many or all of the other components. In such cases fitness is sensitive to initial conditions (Kauffman 1993, 226–7), in the argot of chaologists (call case 1 then “chaos world”). (3) Around \( K = 2 \) a number of high adaptive peaks occur near each other, rather like a mountain range, in a region Kauffman calls the “edge of chaos” (call this “Wright world”). In this region, perturbations in one component, such as mutations, may have little or no effect on other components, as in the ordered regime. Most often the effects will be on a few, but occasionally they can have nearly global consequences (an avalanche, or major rearrangement of relations between components). It is in this region that Kauffman claims that adaptive behavior is possible. The general moral of the story is that natural selection, as an evolutionary agent, is most effective when there are multiple peaks on a fitness landscape.

Wright and Dobzhansky had intuited this, but now Kauffman provides a deep dynamical explanation.

Kauffman applies NK fitness landscapes to provide an explanation of ontogeny and to derive von Baer’s laws. For developing embryos, the NK model predicts that mutations early in development will occur on a relative uncorrelated fitness landscape. Hence, the chances are low that the mutant will be more fit. Late in development, however, the fitness landscape can be assumed to be more correlated. Chances are increasingly better that a mutation will be fit rather than deleterious. Thus, Kauffman’s model generates a redescribed von Baer’s law (Kauffman 1993, 75; cf. Wimsatt 1986). With this notion in place, Kauffman turns to phylogeny by contrasting the Cambrian “explosion” with the Permian “quiescence” (Kauffman 1993, 76–83). Kauffman uses a coevolutionary model of interacting NK adaptive landscapes to argue that the regulatory genetic programs of the newly evolved multicellular organisms in the early Cambrian were still on a rugged landscape, where there was more chance for innovation and for large-scale avalanches in regulatory genetic reorganization, resulting in a disparity of body plans. As the ontogenetic programs of individual species evolved to entrench the earlier phases of development, the later phases would have less rugged landscapes. By the time the Permian extinction and recovery took place, the genetic regulatory programs for multicellular organisms were probably stabilized at or near the edge of chaos. Thereafter, change would have occurred by way of small alterations
in the latter parts of the ontogenetic program of fairly secure developmental regimes, resulting in a greater diversification of species after the extinction, but little or no increased disparity of body plans (cf. Gould 1989).

The Boolean networks that Kauffman uses to model genetic regulation in ontogenetic programs are NK models in which there are just two states of a component (gene), active or inactive (1 or 0). In Boolean networks there are also Boolean functions, such as “AND” or “OR” that control the response of a component to a set of inputs. Again the focus of such models is relational rather than mechanistic. They allow exploration of the behavior of large ensembles of such systems as regulatory genes. Cellular automata, which Kauffman also uses, are just simplified versions of Boolean networks. Kauffman shows that such networks show regimes of ordered behavior (with point or cycle attractors) and chaotic behavior (with enormous chaotic-type attractors), depending on the number $K$ of interactions allowed between genes. Again there is a region in between the ordered and chaotic regimes where there is complex and adaptive behavior (with a fairly small number of reasonably sized attractors). Kauffman cites the rough correlation across a wide phylogenetic range between the number of expected attractors at $K = 2$, predicted by the model for a genome of given size, and the number of cell types, as one indication that such models have an approximate correlation to biological reality.

Kauffman’s models might be of limited interest, except to specialists, unless they provide new ways of thinking about biological systems in general. At the very least, Kauffman’s models might allow a new type of null hypothesis against which the action of natural selection can be measured — a sure sign of programatic changes within the Darwinian tradition (Depew and Weber 1995). For a NK fitness landscape at $K = 0$ there is a single fitness peak on a smooth, correlated landscape and any mutation will cause the system to climb up or down the optimal surface, just as Fisher averred (Fisher 1930). In a moderately rugged fitness landscape, natural selection will use mutation, and the variation that arises from it, to keep the system on a local optimum. Even if selection is strong and maintains the system on such a peak, the shape of the peak, at least distantly, reflects the generic properties of the ensemble as a whole. If selection is weak, the system will wander through the landscape and mutations will drift to fixation by chance. For really rugged and uncorrelated landscapes, as the value of $K$ approaches $K = N-1$, the system will be determined by the generic ensemble properties, and the action of selection will be overpowered. This latter case, which Kauffman calls a complexity catastrophe, provides the null hypothesis against which the effect of natural selection and other evolutionary “forces” is to be measured. Just as the deployment of the Hardy-Weinberg equilibrium
provided a stable background for measuring the action of natural selection in population genetics for the Modern Evolutionary Synthesis, it is possible that use of complex systems dynamics will prove important by giving a new type of stable expectation for systems as complex as genetically regulated organisms seem to be. With the NK model, it is possible computationally to more realistically model such systems and to treat the generic properites that spontaneously arise as the expected baseline against which selection must make its way. Predictions deducible from lawlike behavior can be ventured, from which vantage point the contingencies of history can be tracked. “The establishment of a null hypothesis of this sort is a major accomplishment...” write Burian and Richardson. “Deviations from observed genomic architectures... could be used to detect the pertubing effects of selection and other ‘agents’ of evolutionary change” (Burian and Richardson 1991, 269).

Personally I think that the greatest contribution of Kauffman’s models is to the problem of the origin of life. The NK model allows for a description of a sequence space of catalytic polypeptides (or polyribonucleotides) and a fitness value that is measured by approximation of a region of sequence space to a particular catalyzed chemical reaction. This mapping of catalyzed chemical reactions Kauffman calls a catalytic task space. The map of the estimated hundred million or so possible reactions that could be catalyzed will be covered, though loosely and with very low fitness, by an ensemble of catalytic polymers with random sequence as might be produced abiotically. The likelihood of amino acids and polypeptides forming spontaneously under the putative primitive conditions of the earth has been amply demonstrated (Miller 1953; Fox 1965). The possibility of prebiotic, catalytic polyribonucleotides is also now widely assumed (Gesteland and Atkins 1993). Such an ensemble of catalytic polymers (whether they are polypeptides only, polyribonucleotides only, or a mixed ensemble of both) not only can catalyze a number of chemical transformations, but can also catalyze reactions that lead to making more of the polymers. That is, they become an autocatalytic set. Such autocatalytic systems contain nonlinear terms in the differential equations that describe their chemical kinetics. At the same time, however, they can and do, in computer simulation, show properties of self-organization that characterize actual chemical systems, such as the Belousov-Zhabotinskii reaction (Field and Györgyi 1993). One advantage of Kauffman’s NK models for modeling self-organizing systems is that they require that fewer parameters must be defined than do the differential equations describing chemical kinetics. Another is that they allow an exploration of the potential of the system considered generically and as a whole. In the NK simulation, the greater chemical articulation that develops over time can lead to what Kauffman calls “catalytic closure,” where a connected web of chemical reactions arises in
which every member of the autocatalytic set has at least one of the last steps of its formation catalyzed by some other member of the set. Concomitantly, there arise connected sequences of catalyzed reactions leading from the food set of molecules to all members of the autocatalytic set; that is, in effect, a protometabolism emerges (Kauffman 1993, 298–312).

Kauffman rightly does not envision that this catalytic closure occurs in dilute solution, as has often been assumed in speculations on the origin of life, but rather in some form of phase-separated space, such as within a vesicle (Kauffman 1993, 390–393). Such vesicles would provide not only phase separation but the possibility of concentration of chemicals and catalysts and regulation through control of what did or did not get transported across the membrane of such vesicles, through chemiosmotic processes (Mitchell 1991). By using models based upon random grammars acting upon random strings, Kauffman can simulate the “phase transition” that occurs upon catalytic closure from mathematically finite subcritical sets of catalytic polymers to mathematically infinite supracritical sets of autocatalytic polymers. Even if individual vesicles were slightly subcritical, the ensemble of vesicles as a whole turns out to be supracritical (Kauffman 1993, 404). What is most remarkable about Kauffman’s perspective is that it is not necessary to assume a prior or even concurrent emergence of replicating nucleic acids for there to be a cellular-metabolic process of replication-variation-selection:

Molecular systems, in principle, can both reproduce and evolve without having a genome in the familiar sense of a template-replicating molecular species. It is no small conclusion that heritable variation and hence adaptive evolution, can occur in a self-reproducing system lacking a genome. Since Darwin’s theory of evolution, Mendel’s discovery of the “atoms” of heredity, and Weismann’s theory of the germ plasm, biologists have argued that evolution requires a genome. False, I claim (Kauffman 1993, 285).

Provocative words are these. But they emphasize an important conceptual shift for thinking about the emergence of life and its on-going evolution. From this vanatage point, life emerged out of a sequestered, non-directed, but self-organizing system. This system could make possible the chemical conditions for the synthesis of the ribonucleotides needed for RNA. There are some chemical concerns about the possibility of an RNA-only world, because it is not clear that clays and other non-polymeric catalysts could perform the preferential synthesis of ribose over other sugars and the attachment of purine and pyrimidine bases to ribose (Joyce and Orgel 1993; Orgel 1992). These reactions would not be problematic, however, if the relevant catalytic task space were covered by the ensemble of catalytic peptides. With the presence of ribonucleotides there are reasonable paths, with or without the assistance
of catalytic polypeptides, to RNA capable of catalysis, and more importantly, replication. It follows that, with the advent of information-bearing, replicating nucleic acids in this (proto)metabolic context, the process of selection of the physically stable and chemically efficient would have been enhanced, not created, through the “memory” and “direction” that they would afford those cells processing them. Further, what would have emerged, the biological or natural selection of the reproductively fit, in Lewontin’s sense, emerges as a distinct phenomenon with this development (Weber and Depew 1996). By putting cellular metabolism prior to, or at least co-emergent with, genetic replication, Kauffman is emphasizing a holistic view of living systems that is fully naturalistic, yet non-reductive, a view that sees life as the expected emergent property of deep physical and chemical principles under certain conditions: “We can think of life as an expected emergent collective property of a modestly complex mixture of catalytic polymers” (Kauffman 1993, xvi, original emphasis); and we can conclude that “... the routes to life in the universe are broader than imagined” (Kauffman 1993, 330).

Kauffman’s application of his dynamical models to the origin of life generates suggestions that might have major biological significance. The question that must be asked, in this connection, is how relevant are the implications of these models to biochemical reality? The idea of a proto-metabolism emerging within a vesicle prior to replicating nucleic acids has gained empirical plausibility in recent years. It has been shown that there are sources of amphphilic molecules that will organize into bilayered membranous vesicles, in which plausibly available dye molecules would partition. Such dye molecules could capture light energy and transduce it into osmotic and chemical work (Deamer and Pahley 1989; Deamer and Harang 1990; Deamer 1992; Morowitz 1992). There are, in fact, compelling biochemical and thermodynamic arguments for a cell-first scenario for the origin of life, which are quite independent of Kauffman’s line of reasoning, as has been recently reviewed by Harold Morowitz (Morowitz 1992). Kauffman’s results give additional salience to this body of work.

At a very different hierarchical level, Kauffman applies his models plausibly to simulate the dynamics of ontogenetic programs involved in the development of complex organisms. The models are not per se biological, but they can be set up by analogy to show how such genetic regulatory systems may behave. Chaotic behavior is observed for large values of K. This persists down to about K = 3. At K = 2, however, there is a dramatic emergence of order and a relatively low number of compact state cycles or attractors. For the human genome, assuming N = 100,000 and K = 2 there would be 317 such attractors, which corresponds reasonably well (indeed better than we would have expected) to the observed number of cell types, 254, in humans.
it is far from proving the case, this type of result suggests the possibility that cell types might be defined by attractors in the space of the ensemble of regulatory genes, and that these cell types, as types, exist due to principles of self-organization without natural selection having to do the work of shaping them. “Clearly, if much of the order we see in ontogeny reflects the natural features of complex control systems, we must rethink evolutionary biology. Some of the sources of order lie outside selection” (Kauffman 1993, 408).

This does not deny, however, that selection has an important role to play. “Selection, I suggest, has molded but was not compelled to invent the native coherence of ontogeny” (Kauffman 1993, 410). Such an approach can lead to placing less emphasis on the genealogy of descent and more on shared developmental patterns, a case that has been argued by the developmentalists (Salthe 1993; Goodwin 1994). Kauffman argues that “… we may need selection to account for the particular ensemble selection is exploring but not to account for the fact that organisms remain typical members of that ensemble. Here we confront a new pattern of evolutionary inference: generic properties shared among organisms due to common membership in an ensemble, not by virtue of common descent” (Kauffman 1993, 427). Kauffman goes on to argue that the properties of such ensembles emerge as “macroscopic observables” in a new kind of statistical mechanics. At this point, this claim is a hope rather than a promise, let alone an articulated program. So far forth, these remarks might suggest that Kauffman has been pulled into the developmental tradition. But Kauffman pulls back from leaping away from the Darwinian tradition. Ultimately he sees selection as a co-creating partner of self-organization by proposing a new and deeper role for natural selection as a phenomenon than Darwinians have thus far envisioned. It is selection itself that pulls systems into the region of adaptive behavior that emerges at or near the interface of the ordered and chaotic regimes:

If it proves true that selection tunes genomic systems to the edge of chaos, then evolution is persistently exploring networks constrained to this fascinating ensemble of dynamical systems. The generic properties of this perhaps most important ensemble emerge as the best hypothesis to account for the remarkable order in organisms. If this overall view proves useful, then the manifold marriage of self-organization and selection consists in constraint of genomic systems to this well-wrought ensemble as a result of selection’s achieving systems best able to adapt … (Kauffman 1993, 522).

The action of natural selection is to favor those entities that the interaction of self-organization and selection have produced and to maintain such systems in the region of phase space that affords them the greatest opportunity of further evolution. Thus adaptability is itself an adaptation. This gives a dynamical
backing to the notion that evolution by natural selection is most effective when it can, in some manner, store variation for changing environments rather than merely using it up. As such, Kauffman can be viewed as developing the Dobzhanskyan research program. Further, John Campbell and Christopher Wills have argued, upon quite different premises, that natural selection will favor traits that enhance the possibility of further evolution, and so evolvability can be seen as the greatest adaptation of all (Campbell 1987; Wills 1989). As abstract as Kauffman’s dynamical models may be, they do seem able to capture in a general way important biological phenomena.

Of course Kauffman’s models are an oversimplification. They can, and will, be criticized by practitioners of the various disciplines to which he has applied them. It is not remarkable that they are imperfect. What is remarkable is that they work as well as they do. They demonstrate the potential of employing models of complex systems dynamics to explore the origin and evolution of living beings and to suggest new types of theoretical and experimental exploration. The usefulness of Kauffman’s models is that they are level independent and do not include any causality beyond how elements of an ensemble can be related. Kauffman rarely addresses the issue of underlying causality, but when he does, he acknowledges the fact that real systems that show the type of dynamics he is exploring are maintained far from equilibrium by energy flows (Kauffman 1993, 389, 393). Having made this admission, however, he does not make any use of the tools of nonequilibrium thermodynamics or kinetics that are available (Nicolis and Prigogine 1977, 1989; Harrison 1988; Brooks and Wiley 1986, 1988; Wicken 1987; Swenson 1989, 1996, for example). What Kauffman is interested in is the dynamical consequence for complex systems, the components of which are related in specific ways, under these types of thermodynamic constraints. For this, the NK-ensemble, fitness-landscape models and Boolean networks serve well enough. Whatever its limitations, the main significance of Kauffman’s work is that he has shown the possibility that many phenomena of contemporary evolutionary science flow rather easily and directly from assumptions taken from complex systems dynamics. What Kauffman’s models allow us to do is to situate biological complexity against a dynamical background that renders it expectable, tractable, and comprehensible.

References


