Emergence of Order in Small Autocatalytic Sets Maintained Far from Equilibrium: Application of a Probabilistic Receptor Affinity Distribution (RAD) Model

Doron Lancet, Ora Kedem, and Yizhaq Pilpel

Department of Membrane Research and Biophysics, The Weizmann Institute of Science, Rehovot 76100, Israel

Key Words: Catalysis / Computer Experiments / Light, Absorption / Molecular Interactions / Nonequilibrium Phenomena

We examined the behavior of auto-catalytic sets of polymers by a computer simulation. Polymers are allowed to interact with each other, whereby each polymer molecule may catalyze the formation and degradation of others. The system is subjected to a set of thermodynamic and kinetic constraints, including a constant influx of free energy, which keeps the system away from chemical equilibrium and thus enables the effect of catalysis. The system is found to continuously change and probe many possible values in the composition space. In this simulation we make use of a Receptor Affinity Distribution (RAD) model to predict the probabilities of interaction and catalysis. Our results indicate that initially random sets of polymers, under the assumptions of the model, might accumulate information (i.e., clustering in the composition space). Sets will occupy a limited region of composition space, and temporarily reproduce themselves or disperse and give rise to other sets.

Introduction

One of the major open questions related to the origin of life is how primordial monomers (e.g. amino acids and nucleotides) spontaneously assembled into information-rich oligomers and polymers with specific sequence, capable of carrying out catalysis and replication. One set of recent models [1–4] proposes that this happened through the emergence of ensembles of molecules capable of mutual catalysis. These may display a capacity of self-replication of the ensemble *as a whole*, although none of the individual molecular components is necessarily self-replicating. We propose here that such models may be further explored through the application of a statistical metric for the probability of mutual catalysis for biopolymer formation and degradation.

The RAD Model

The Receptor Affinity Distribution or RAD model [6, 7] was developed originally in order to answer the following question: given a large number (N) of different receptors of any kind (R_1, $R_2 \ldots R_N$) and an arbitrary ligand (L), what is the affinity between the ligand L and each of the receptors R_i. Or more generally, what is the probability distribution of such affinities. There are two basic assumptions in this model. First, the binding sites of both the receptors and the ligands are regarded as a mosaic of potential elementary interactions (e.g., ionic, dipole, hydrophobic, H-bond). Second, in its simplest form, the model assumes that the contribution of each elementary interaction to the total free energy of the reaction is equal. The model provides a phenomenological way of computing the extent to which members in this mosaic complement each other (cf. Fig. 1). The number of "successful" elementary interactions (M) is distributed binomially. The actual value of M allows the calculation of the free energy of the reaction. Consequently, it is possible to derive a distribution $\Psi(K)$ of the binding constants K_i. Despite its simple assumptions, the model is capable of making predictions, including the size of the repertoire of the olfactory and the immune system [6, 7].

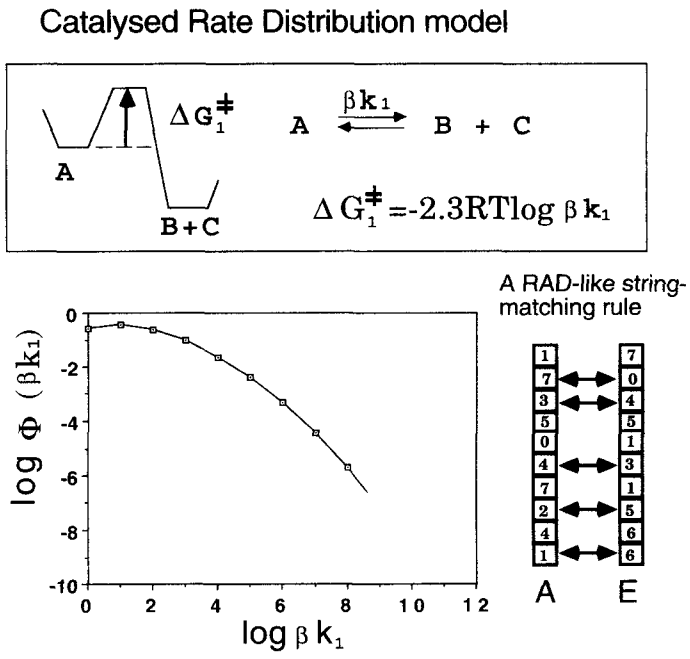

Fig. 1
A Catalyzed Rate Distribution model – an extention of the Receptor Affinity Distribution (RAD) model to reaction kinetics. For example, the probability distribution $\Phi(\beta k_1)$ of a forward reaction rate k_1 catalyzed by a catalyst E to an extent β may be derived using a "string matching rule". $\Phi(\beta k_1)$ is assumed to be a continuous distribution since $\Delta G^{\neq}$, the free energy of activation, may assume any positive value. As in the RAD model, the reactant A as well as the catalyst E are represented by strings of integers, and the catalytic site complementarity is symbolically depicted as "subsites" whose digits sum up to 7. Catalytic potency is assumed to be related to the number of "successful interaction" (5 in the example shown). Parallel to the RAD model, $\Phi(\beta k_1)$ may be a binomial distribution (cf. ref. 7)

 © VCH Verlagsgesellschaft mbH, D-69451 Weinheim, 1994 0005-9021/94/0909-1166 $ 10.00+.25/0

Extending the RAD Model: Application to Catalysis

To analyze catalytic sets and predict the probability of catalytic potencies we propose here to extend the RAD model. Consider the isomerization reaction,

$$A \overset{E}{\longleftrightarrow} B$$

where E serves as a catalyst. For the uncatalysed reaction, let k_1, k_{-1} be the rate constants for forward and backward reactions, respectively. Let $\beta \cdot k_1$, $\beta \cdot k_{-1}$ be the respective catalyzed rate for the above reaction where β is the *catalytic value*, a multiplicative factor (equal for both directions) which signifies the extent of catalysis; $\beta = 1$ indicates the absence of catalysis and $\beta > 1$ indicates its presence. Typical values for non-enzymatic catalysis range from $\beta = 10$ to $\beta = 100$, while for enzymes, values as high as $\beta = 10^6$ are observed.

Applying the RAD concept to catalysis amounts to developing a distribution function for the likelihood that a particular value of rate enhancement β will occur when considering an arbitrary substrate and a large set of potential catalysts. Is it justified to assume that such a distribution resembles the Receptor Affinity Distribution?

In order to get $\beta > 1$ for the above reaction, the catalyst E has to bind to the potential substrate A with a minimal affinity. A catalyst-reactant complex is included in practically all mechanistic descriptions of catalysis. In the case of enzyme catalysis, a specific multi-site, non-covalent attachment between the enzyme and its substrate is assumed to be an important determinant of catalytic efficiency. This is true also for more complex reaction schemes, such as bimolecular condensations, whereby both reactants are assumed to bind to the catalyst simultaneously. Therefore, if binding *per se* obeys an affinity distribution law, catalysis (i.e. β) may behave similarly.

Catalytic antibodies [8] provide an interesting corroboration of the concept of "probabilistic catalysis". The underlying notion is that in a large collection of biopolymers (such as the antibody repertoire), which have not specifically evolved to serve as enzymes, there exist members which can catalyze (with different values of β) an arbitrarily chosen reaction. This is consistent with the notion that a distribution of β values may be defined for a substrate among a large number of potential catalysts. In other words, a RAD-like model should be applicable to the problem of catalytic values distribution (Fig. 1).

Recall that the RAD model refers to the binding values in a graded manner, as the binding constant (K) may assume any real value, within chemically meaningful limitations. Likewise, $k_{\pm 1}$, and therefore β, may assume any value within a continuous range. This is justified since the effect of catalysts is to lower the activation-free energy of the reaction, and since β obeys the relation:

$$\Delta(\Delta G^{\neq}) = -RT \ln \beta$$

it is justified to claim that the catalytic values change in a graded manner.

What is the significance of such distribution? Kauffman [3, 4] in his model describing the behavior of auto-catalytic sets of polymer, assigns a fixed probability P for each polymer to catalyze any reaction, noting that the value of P might differ from one polymer to another. In order to calculate the variation with time of individual polymer components in our system, it is necessary to assign actual kinetic constants to the reactions involved. For this, it may be useful to derive the extent of catalysis β for each pair of components based on a RAD-like model (Fig. 1).

The RAD model, as well as the RAD-like model for catalysis, clearly constitute over simplifications of complexity of molecular interactions among biopolymers. But since rigorous predictive models for molecular docking and catalysis are not available, we propose here to implement the RAD model in ways that may capture some of the complexity that governs the reactions in an autocatalytic set.

Emergence of Order in Small Autocatalytic Sets of Biopolymers

These concepts are the basis of a computer simulation in a system composed of a set of polymers (symbolically represented in the computer as strings of characters). Polymers are allowed to interact with each other and catalyze the formation and degradation of each other. The probability of these events is computed with a simple binding metric, according to a RAD-like model. We wish to explore the potential of such system to gradually accumulate order and information. Specifically, to see whether it is likely to start a run of the simulation program with a random set of polymers, such that the average probability of mutual catalysis is low, and observe the spontaneous appearance of a collection of polymers that constitute an auto-catalytic set. This approach is akin to the analysis of "self programmable matter" by Rasmussen and colleagues [5].

The simulated model is designed such that it obeys a set of thermodynamic and kinetic constraints. We assumed a constant influx of free energy (supplied, for example, by light). The consequence is that the system is maintained far from equilibrium. Obviously, we cannot expect any changes due to catalysis in the concentrations of polymer species at equilibrium. Furthermore (see Fig. 2), it is possible to derive an expression relating the degree of deviation from equilibrium (or the driving force of the reaction) to the effect of a catalyst on the net reaction rate. This relation shows that the further the system is kept away from its equilibrium (in either direction), the bigger is the effect of catalysis.

In the simulation we update the composition of a collection of polymers at each time cycle. This is done by choosing a polymer as catalyst and monomers as potential substrates. Making use of the RAD-like model, we predict the probability of the reaction that creates a new polymer or degrades an existing one. In addition, at every time cycle, spontaneous formation or degradation may occur with a small pre-assigned probability.

The main tool to view the results of the simulation is to follow the evolution, with time, of an index (H) which

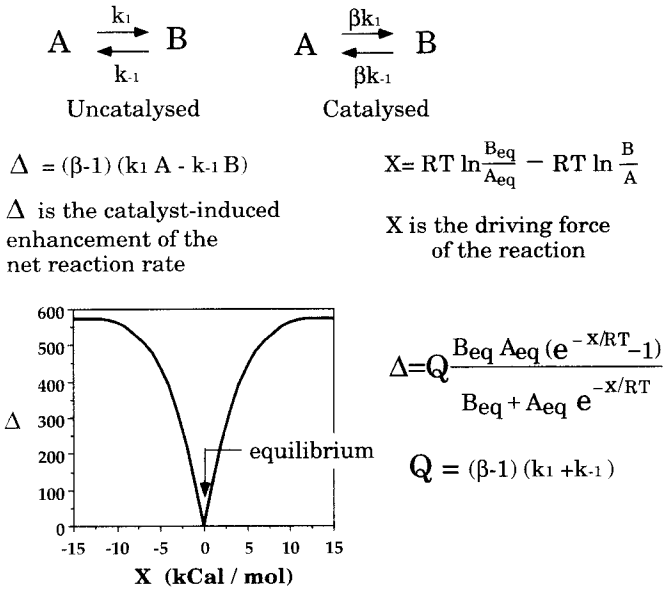

Fig. 2
The dependence of the extent of catalysis (Δ) on the driving force of the reaction (X), related to the deviation from equilibrium. Δ, shown in *absolute value*, has units of rate (M/s). It is the increase, due to catalysis, of the net reaction rate (the difference between the forward and backward reaction). β is the catalytic value, a multiplicative factor (equal for both directions) which signifies the extent of catalysis. Note the difference between "catalytic value" β, which is an inherent property of a catalyst-substrate pair, and "extent of catalysis" Δ, which changes as the reaction proceeds towards equilibrium and thus depends on actual concentrations. Δ is 0 when the driving force is 0, i.e. at equilibrium. It rises steeply over a range of a few kcal/mol, then reaches a plateau whose value depends on the rate constants and the concentrations.

In the example shown the catalytic value $\beta = 20$, the rate constants are $k_1 = k_{-1} = 0.2\,s^{-1}$ and the equilibrium concentrations are $A_{eq} = B_{eq} = 75$ mM. R is the gas constant and T the absolute temperature (300 K)

measures "clustering" in the composition of the potentially auto-catalytic set. The index of clustering is given by:

$$H = \sum_i \sum_j 1/X_{ij}$$

where X_{ij} is the vectorial distance between the vectors representing the polymers i, j. This is similar, in principle, to a previously described mutual information measure (Knudsen et al. 1991).

The results of a preliminary simulation are shown in Fig. 3. It may be seen that the order parameter H displays an erratic behavior, with spontaneous increases and declines in the range between 0 and 150. The behavior is very sensitive to initial conditions, and in this it is not unlike some systems manifesting chaos. The peaks in the graph represent polymer sets with high degree of mutual similarity, but adjacent peaks may correspond to completely different set compositions.

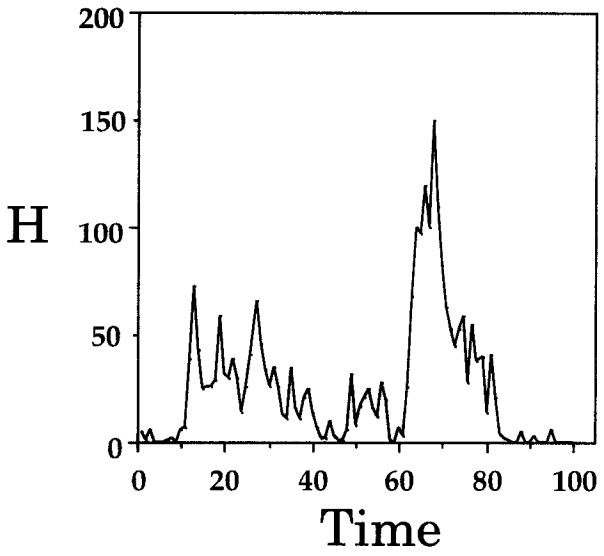

Fig. 3
The results of a computer simulation in which the behavior of a set of trimers, composed out of 10 monomer types, is followed over time. H is an order parameter as defined in the text. At each time cycle, a candidate trimer is chosen for synthesis, and its fate is decided based on a single random catalytic encounter with one of the existing trimers. The extent of catalysis, hence the probability of synthesis, is decided based on a Catalyst Rate Distribution formalism (see Fig. 2). We assign $\beta = 1$ (no catalysis) for all but the top 5% on the RAD-like metric scale, and a linearly increasing β within the top interval. The simulation also includes spontaneous synthesis and decomposition. The system is continuously maintained away from equilibrium by a "forced decomposition" reaction that could symbolize light-induced depolymerization. H is seen to vary erratically, with occasional peaks of up to $H = 150$. In a control simulation with no catalysis the average was $H = 5.8 \pm 5.6$

Conclusion

The approach presented here shows the potential applicability of a model for ligand-receptor interaction to the analysis of mutually catalytic biopolymers in autocatalytic sets. Considerable added effort will be needed to analyze the simulated behavior in detail, and establish its potential relationship to very early stages of transition from random polymerization of monomers, such as amino acids and nucleotides, to information-rich biopolymers.

References

[1] F.J. Dyson, Origins of life, Cambridge University Press, Cambridge 1985.
[2] F.J. Dyson, A model for the origin of life, J. Mol. Evol. *18*, 344 (1982).
[3] J.D. Farmer, S.A. Kauffman, and N.H. Packard, Autocatalytic replication of polymers, Physica 22D, 50 (1986).
[4] S.A. Kauffman, The origins of order – Self-organization and selection in evolution, Oxford University Press, Oxford 1993.
[5] C. Knudsen, R. Feldberg, and S. Rasmussen, Information dynamics of self-programmable matter, in: Complexity, Chaos and Biological Evolution, ed. by Mosekilde and Mosekilde, Plenum, New York, pp. 223–245 (1991).

[6] D. Lancet, A. Horovitz, and E. Katchalski-Katzir, Molecular recognition in biology: Models for analysis of protein/ligand interactions. In: Perspectives in Supramolecular Chemistry, ed. by Behr, Wiley, New York 1994 (in press).
[7] D. Lancet, E. Sadovsky, and E. Seidemann, Probability model for molecular recognition in repertoires of biological receptors: significance to the olfactory system, Proc. Natl. Acad. Sci. *90*, 3715 (1993).
[8] R.A. Lerner, S.J. Benkovic, P.G. Schultz, At the Crossroads of chemistry and immunology: catalytic antibodies. Science *252*, 659 (1991).

Presented at the Joint Meeting and International Bunsen Discussion Meeting of the Deutsche Bunsen-Gesellschaft für Physikalische Chemie "Self-Organization of Biopolymers" in Jena, April 10th to 13th, 1994 E 8750