

in Bak, Tang, and Wiesenfeld's "self-organized criticality" (9). In an essay entitled "More Is Different," Anderson (10) described how features of organization may arise as an "emergent" property of systems. An example of this point of view is given by work on complexity "phase transitions" and accompanying speculations that various aspects of biological systems sit on a critical point between order and complexity (11).

The next few years are likely to lead to an increasing study of complexity in the context of statistical dynamics, with a view to better understanding physical, economic, social, and especially biological systems. It will be an exciting time. As science turns to complexity, one must realize that complexity demands attitudes quite different from those heretofore common in physics. Up to now, physicists looked for fundamental laws true for all times and all

places. But each complex system is different; apparently there are no general laws for complexity. Instead, one must reach for "lessons" that might, with insight and understanding, be learned in one system and applied to another. Maybe physics studies will become more like human experience.

References and Notes

1. U. Frisch, B. Hasslacher, Y. Pomeau, *Phys. Rev. Lett.* **56**, 1505 (1986); J. Hardy, O. de Pazzis, U. Frisch, *J. Math. Phys.* **14**, 1746 (1973); *Phys. Rev. A* **13**, 1949 (1976).
2. Early work on the derivation of hydrodynamics from conservation laws can be found in S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge Univ. Press, Cambridge, ed. 3, 1970).
3. A. R. Kerstein, *J. Fluid Mech.* **291**, 261 (1997); S. Wunsch, thesis, University of Chicago (1998). For experiments, see, for example, B. Castaing et al., *J. Fluid Mech.* **204**, 1 (1989). For theory, see E. Siggia and B. Shraiman, *Phys. Rev. E* **49**, 2912 (1994).
4. B. Mandelbrot, *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk* (Springer-Verlag, New York, 1997).
5. A. Katchalsky and P. F. Curran, *Nonequilibrium Processes in Biophysics* (Harvard Univ. Press, Cambridge, MA, 1967).
6. G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems* (Wiley, New York, 1977).
7. A. Turing, *Philos. Trans. R. Soc. London Ser. B* **327**, 37 (1952).
8. For example, L. Kadanoff, A. Libchaber, E. Moses, and G. Zocchi [*Recherche* **22**, 629 (1991)] discussed the development of interlinked structures in a Rayleigh-Benard flow.
9. P. Bak, C. Tang, K. Wiesenfeld, *Phys. Rev. Lett.* **59**, 381 (1987); J. M. Carlson, J. T. Chayes, E. R. Grannan, G. H. Swindle, *ibid.* **65**, 2547 (1990).
10. P. W. Anderson, *Science* **177**, 393 (1972).
11. S. A. Kauffman, *The Origin of Order* (Oxford Univ. Press, Oxford, 1993); *At Home in the Universe* (Oxford Univ. Press, Oxford, 1995).
12. Supported in part by NSF grant NSF-DMR-93-14938 (N.G.) and by the ASCI Flash Center at the University of Chicago under U.S. Department of Energy contract B341495 (L.P.K.).

VIEWPOINT

Complexity in Chemistry

George M. Whitesides* and Rustem F. Ismagilov

"Complexity" is a subject that is beginning to be important in chemistry. Historically, chemistry has emphasized the approximation of complex nonlinear processes by simpler linear ones. Complexity is becoming a profitable approach to a wide range of problems, especially the understanding of life.

"Complexity" is a word rich with ambiguity and highly dependent on context (1). Chemistry has its own understandings of this word. In one characterization, a complex system is one whose evolution is very sensitive to initial conditions or to small perturbations, one in which the number of independent interacting components is large, or one in which there are multiple pathways by which the system can evolve. Analytical descriptions of such systems typically require nonlinear differential equations. A second characterization is more informal; that is, the system is "complicated" by some subjective judgment and is not amenable to exact description, analytical or otherwise.

In chemistry, almost everything of interest is complex by one or both definitions. Consider the design and synthesis of a simple organic substance ($<10^2$ covalently bonded, first-row atoms) as a candidate drug—a representative activity for organic, medicinal, and biological chemists. A single step in the multistep synthesis of such a substance might involve 10^{22} molecules of several types (each

comprising as many as 10^2 anharmonically oscillating bonds) and several times this number of interacting nuclei and electrons, all immersed in 10^{24} molecules of solvent. The synthesis itself might proceed by perhaps 10 different strategies (that is, sequences of reactions) for making and breaking bonds and for generating the intermediate compounds that ultimately result in the final compound; each strategy might have many thousands of possible variants differing in synthetic detail. The design of a molecule that has the right properties (shape, surface properties, and associated electrostatic fields) to interact specifically with one part of the surface of a target protein molecule presents yet another set of complicated challenges (Fig. 1) (2).

Faced with the impossibility of handling any such real system exactly, chemistry has evolved a series of approaches to the treatment of complex systems, which range from reasoning by analogy, through averaging, linearization, drastic approximation, and pure empiricism, to detailed analytical solution. The study of complexity in systems of reactions (or of processes or of properties) that can be described by nonlinear equations has been limited to the few that are both complex enough to be interesting and simple enough to be tractable. The emphasis in thinking

about complicated systems has been to find methods that are predictive, even if they are nonanalytical. Philosophically, chemistry is a branch of science that attempts to predict and control rather than simply to observe and analyze: A large industrial reactor that produces heat in unpredictable bursts is more immediately terrifying than interesting. The optimization of combustion for the production of work, the understanding of mechanisms of drug action, and the development of strategies for organic synthesis are all problems of great complexity. They are also problems of sufficient urgency, which must be solved as best as possible, even if analytical solutions for them are not practical.

Chemistry is now evolving away from the manipulation of sets of individual molecules and toward the description and manipulation of systems of molecules, that is, living cells and materials. This evolution toward complexity is, perhaps counterintuitively, generating new types of problems that are sufficiently simple in some aspects for "complexity" in its analytical sense to provide a valuable way of thinking about them. These problems are often at the border between chemistry and other fields such as physics, biology, biophysics, and materials science. They may represent efforts to describe properties (for example, flux through a catalytic pathway in metabolism, distribution of greenhouse gases in the atmosphere, and fracture toughness of a polymer) that strongly depend on time, space, and conditions and in which the granularity of the description that is de-

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA.

*To whom correspondence should be addressed. E-mail: gwhitesides@gmwgroup.harvard.edu

sired is relatively low (even if the molecular complexity is intractably high).

The objective of this survey is to outline some of the strategies that have been developed in chemistry to deal with complex problems. A subset of these problems (for example, oscillating reactions) provides classical examples of complex systems in the sense that they can be described analytically by relatively simple sets of nonlinear differential equations. Other subsets (for example, the definition and evaluation of the relative complexity of different synthetic pathways) have been treated analytically with partial success; here, empirical or analog approaches may guide the construction of more analytical strategies (3). Yet others (for example, recognizing complementarity in molecular shapes) represent problems of general importance for which there are no simple general solutions (4). Even the definition of "shape" is difficult because many molecules are dynamic and exist in a manifold of rapidly interconverting conformations (5). Because complexity is ubiquitous throughout chemistry, the selection of the subjects discussed is necessarily both personal and incomplete.

Linearization and Reasoning by Analogy

The approaches to complicated problems that have historically been most favored in chemistry are simplification, linearization, and extensive reliance on reasoning by analogy. The study of the mechanisms by which reactions occur usually begins by establishing the dependence of these rates on the temperature and concentrations of reactants and products. Processes occurring by sequences of chemi-

cal reactions occur in a very wide range of complexities, from very simple (the combination of two H atoms in the formation of H₂) to very complex (explosions, metabolic cycles in cells, and systems of reactions that are important in environmental chemistry). Simple one-step reactions in solution can often be described, with the transition-state theory, by an exponential equation that relates the rate constant k to the difference in free energy ΔG^\ddagger between the transition and ground states (6)

$$k = \frac{k_b T}{h} e^{-\Delta G^\ddagger/RT} \quad (1)$$

$$\ln k = -\Delta G^\ddagger/RT + C \quad (2)$$

where k_b is Boltzmann's constant, T is temperature, h is Planck's constant, R is the gas constant, and C is a proportionality constant. The fact that even fairly complicated variants of this type of equation can often be linearized successfully has concentrated attention on equations that can be linearized.

The emphasis on systems that can be linearized has extended fruitfully to more complex systems. Linearization is, for example, at the heart of a very successful style of analog reasoning, whose many variants are subsumed under the title "physical-organic chemistry" (7). Physical-organic chemistry was developed to provide a semiquantitative description of the mechanisms of the organic reactions that form the backbone of organic synthesis. Even the simplest of these reactions is usually unmanageably difficult if considered rigorously; the reacting species may exist in solution in many interconverting shapes, local and nonlocal variations in structure (so-called substituent and steric effects)

influence reactivity, and the interaction of the reacting species with solvent molecules and local ions is important. The approach taken in physical-organic chemistry is to admit that the microscopic details of the reaction are inaccessible and to adopt a frankly empirical approach based on a comparison of similar reactions. The procedure begins by selecting a reaction (reaction a) for which extensive data that correlate structure with reactivity are already available. If one then wishes to predict the influence of molecular structure on a second reaction (reaction b), which can be argued to be similar to reaction a, one hypothesizes that changes in the structure of reactants in the two types of reactions result in proportional changes in the free energies ΔG^\ddagger that characterize them (Eq. 3 is a so-called linear free-energy relationship).

$$\Delta G_a^\ddagger \propto \Delta G_b^\ddagger \quad (3)$$

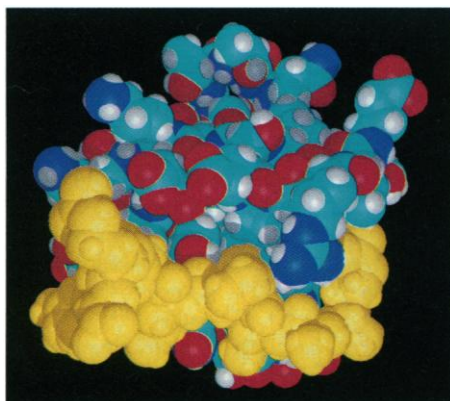
An astonishing degree of predictivity in chemistry can be achieved with these types of linearizations.

Other fundamental problems in chemistry (for example, the description of the shapes of molecules and the estimation of the complexity of a synthetic pathway) have not been amenable to broadly useful nonempirical mathematical formulation and particularly not to formulation in terms of linear equations. These problems are largely left as a subjective art and described in graphical and iconic form (5).

Tractable Complexity

In the sequence of complexity—from static equilibrium, to dynamic steady state, to dynamic complexity, to chaos—there are only a

Fig. 1 (left). Space-filling model of the peptide RLP2 (yellow spheres) bound to the c-Src SH3 domain (red spheres, O atoms; light blue spheres, C atoms; dark blue spheres, N atoms; white spheres, H atoms) (2). This domain is a component motif of proteins making up signaling pathways within the cell. The peptide ligand was discovered by a combinatorial approach. Understanding and controlling cellular signaling is an example of a biochemical problem of high complexity. **Fig. 2 (right).** Oscillating traveling waves in a Belousov-Zhabotinsky reaction. Coupled chemical reactions cause changes in concentrations of the reagents that, in turn, cause local changes in the oxidation potential of the solution. These potentials are visualized with a redox indicator dye. (Photograph used with permission of Felice Frankel, copyright 1997.)



few, carefully chosen sets of chemical reactions whose properties make them appropriate as case studies in complexity. Oscillating reactions of the type represented by the Belousov-Zhabotinsky reaction are perhaps the best known example (8). This class of reactions has the characteristics that the simultaneous operation of two processes, reaction and diffusion, results in a system in which the concentrations of reactants and products oscillate temporally and spatially and in which this oscillation can result in ordered patterns (Fig. 2). These reactions can be described mathematically by systems of nonlinear equations of greater or lesser complexity, but Eqs. 4 and 5 represent a minimum set of two reaction-diffusion equations (9)

$$\partial u/\partial t = F(u, v) + D_u \nabla^2 u \quad (4)$$

$$\partial v/\partial t = \varepsilon G(u, v) + D_v \nabla^2 v \quad (5)$$

Here, u is the concentration of a species that catalyzes reaction; v is the concentration of a species that inhibits reaction; $\partial u/\partial t$ and $\partial v/\partial t$ describe changes in concentrations of u and v , respectively, with time; $F(u, v)$ and $\varepsilon G(u, v)$ characterize reactions between u and v , respectively; and D_u and D_v are the diffusion coefficients of u and v , respectively.

A stunning example of behavior that is oscillatory in time and space and that occurs in a situation approaching the real world is the heterogeneous Pt-catalyzed oxidation of CO to CO₂ by O₂ (10). This reaction proceeds by waves of oxidation of surface-bound CO moieties, followed by desorption of CO₂ and adsorption of CO; these reactions are accompanied by a reconstruction of the Pt surface (11). These reactions form standing waves and spirals and generate transient pulses in concentrations. It is possible to control the form of these patterns by patterning the surface of the catalyst (12).

An important motivation in chemistry for studying complexity has been to learn about processes in living systems. One of the most striking characteristics of cells is the sheer complexity of metabolism (13). The human genome probably has on the order of 10⁵ expressed gene products; many of these proteins are enzymes, receptors, and members of signaling sequences, that is, functional parts of metabolism. Understanding a system with this many interacting components (especially at the current, very incomplete understanding of the functions and interconnections between them) is clearly out of the question. A more tractable problem is to examine discrete, relatively self-contained sections of metabolism. One metabolic cycle that has been studied in substantial detail is glycolysis, that is, the conversion of glucose to pyruvate with the production of adenosine 5'-triphosphate and the reduced form of nicotinamide adenine dinucleotide (14). This se-

quence of reactions involves 10 enzymes, with various levels of modulation of the catalytic activities of some of these enzymes by the products of others. The concentrations of intermediates in the cell in this sequence seem, in some circumstances, to oscillate, and it is possible to rationalize this oscillation with understandable models (15).

Complex processes of interest in chemistry extend, of course, beyond molecules to materials (16). Examples of topics of current interest in chemistry are the properties of glasses (17) and the flow of fluids (in part for its relevance to microanalytical systems) (18). The latter has been examined particularly extensively. At low rates of flow or in small channels (circumstances described by low values of the Reynolds number Re) the flow of liquids is laminar; as rates increase, flow often develops complex regular structures. At even higher rates, flow becomes turbulent. The description of the early stages of this progression from laminar to turbulent flow is well described and has served as one of the early foci of considerations of complexity; the latter stages remain difficult problems in fluid dynamics (19).

Outlook

In considering complicated processes, the most common strategy in chemistry has been to simplify and linearize or to use nonanalytic methods to guide operations. "Complexity" per se—the study of nonlinear processes with high sensitivity to conditions—has not been the focus of major effort. A number of new types of problems are, however, moving in directions that will require better intellectual tools for understanding complex systems. Whether these problems will require or contribute to general theories (20) of complexity remains to be seen.

One motivation is certainly the pervasive opportunity provided by genomics and the opportunities of the proteomics era. It is certain that research over the coming years will generate complete DNA sequences for many organisms, and the problem of how to use those data is now a key concern in biology and biochemistry. The sequence data will unquestionably be most useful if they can be converted to models of organisms. DNA sequences code for the sequences of amino acids in proteins, and one can optimistically expect that efficient methods for establishing the structures and functions of these proteins will be developed. All of this information, however, will not result in a model for the organism unless the operation of the components in the organism, with feedback and with the complex cooperative interactions that characterize biological systems, can be understood. Studies of metabolic cycles and signaling pathways suggest that it will be possible to isolate parts

of metabolism from the whole and to study these parts; these studies also demonstrate the current difficulty in rationalizing even the behavior of these relatively simple systems, much less the emergent properties of organisms.

In part in reaction to the difficulty of this task, chemistry has recently emphasized the development of combinatorial methods, which were originally used for discovering leads for new drugs and more recently for materials and catalysts (21). Combinatorial methods are, in a sense, an approach to complexity that is the antithesis of analytical methods: instead of understanding, modeling, and predicting, one throws dice. Is there a connection between these two approaches? Perhaps. One approach to recognizing patterns hidden in combinatorial data might, in principle, be through statistical analytical methods of the types that are represented by artificial neural nets (ANNs) (22). Although these methods can recognize patterns and can model complex behavior, their use in chemistry has been limited for several reasons, one of which is the fact that much of chemistry is information-poor. ANNs work best when they can be trained with large fields of data that describe both the areas that immediately surround the desired solution and the peripheral areas of less interest; in other words, the net must learn not only the location of the mountain peak, but also the shapes of the surrounding slopes and valleys. Making and testing even a single compound for biological activity has required a substantial effort, and in the past, an extensive program in medicinal chemistry might examine only a few thousand compounds. This relatively small number and the search strategy customarily used to guide the syntheses do not provide the information needed to train ANNs to recognize patterns as complex as those that make up the biochemical pathways being targeted by drugs. Combinatorial methods, by making it possible to consider libraries of billions of compounds, may (or may not) be able to provide the quantity and type of information required for the effective use of ANNs and other statistical pattern recognition schemes as tools for examining complexity.

Additional understanding of complexity in chemical systems may also be developed by examining the behavior of very simple systems—single molecules (23). Chemistry has relied heavily on the ability of ensemble properties that are obtained through thermodynamics and statistical mechanics to make it unnecessary to consider the behavior of individual molecules. Single-molecule chemistry is, however, now making it possible to inquire about the variety of individual molecular behaviors. Understanding how the properties of single molecules aggregate into the familiar averaged properties of macroscopic samples of chemicals will help to tease apart

the threads of complexity in chemical systems. The promise of this new activity is particularly rich for macromolecules (including biologically relevant macromolecules), in which opportunities for the existence of many different molecular conformations, each with different properties, are high.

At the core of chemical interest in complexity are the two fundamental problems concerning life, that is, trying to understand (i) how collections of molecules can give rise to the varieties of behaviors that characterize cells and organisms and (ii) how individual molecules might have originally assembled into collections that had the characteristics of life (energy dissipation, self-replication, and adaptation). Whether the understanding of complexity at the molecular level will reveal important elements of the structure of life is unclear. We do not know if it is conceptually possible to connect molecular-level processes to organismic behavior deterministically. Certainly, knowing everything about the electronic properties of Si and the operating characteristics of transistors tells very little about the higher level characteristics of computers.

Fortunately, there is also the inverse opportunity: learning from biological complexity as a method of stimulating new chemistry. With this opportunity, there is great reason for optimism. Biological systems display such a large number of remarkable capabilities (and capabilities that are so clearly complex) that their analysis will unquestionably be a rich source of models for new areas of

chemistry. ANNs are one example of a successful transfer of information about a complex biological system to nonbiological applications. ANNs were developed, in part, as a tool with which to model the brain. To what extent current ANNs do so is a continuing subject of discussion, but the effort to make the connection between ANNs and brains (and to learn from the brain) has unquestionably expanded the capabilities of computation. In this same sense, biology (and perhaps also complex materials) offers examples of complex systems that show types of behavior that are now uncommon in molecular chemistry. One of the opportunities in fundamental chemical research is to learn from biology and to use what is learned to design nonbiological systems that dissipate energy, replicate, and adapt. Whether such systems would model life is moot; they would unquestionably be very interesting and probably very important.

References and Notes

1. K. Mainzer, *Thinking in Complexity* (Springer-Verlag, New York, ed. 3, 1997).
2. S. Feng, J. K. Chen, H. Yu, J. A. Simon, S. L. Schreiber, *Science* **266**, 1241 (1994). Atomic coordinates were obtained from the Protein Data Bank [J. L. Sussman et al., *Acta Crystallogr.* **D54**, 1078 (1998)].
3. J. Hendrickson, *Top. Curr. Chem.* **62**, 49 (1976); H. W. Whitlock, *J. Org. Chem.* **63**, 7982 (1998).
4. P. G. Mezey, *Int. Rev. Phys. Chem.* **16**, 361 (1997).
5. R. W. Hoffman, *Angew. Chem. Int. Ed. Engl.* **31**, 1124 (1992).
6. J. I. Steinfeld, J. S. Francisco, W. L. Hase, *Chemical Kinetics and Dynamics* (Prentice-Hall, Englewood Cliffs, NJ, 1989).
7. F. A. Carroll, *Perspectives on Structure and Mechanism in Organic Chemistry* (Brooks/Cole, Pacific Grove, CA, 1998).
8. I. R. Epstein and K. Showalter, *J. Phys. Chem.* **100**, 13132 (1996); A. N. Zaikin and A. M. Zhabotinsky, *Nature* **225**, 535 (1970).
9. V. S. Zikov, A. S. Mikhailov, S. C. Muller, *Phys. Rev. Lett.* **81**, 2811 (1998).
10. R. Imbibi and G. Ertl, *Chem. Rev.* **95**, 697 (1995).
11. G. Ertl, *Science* **254**, 1750 (1991).
12. M. D. Graham et al., *ibid.* **264**, 80 (1994).
13. B. Alberts et al., *Molecular Biology of the Cell* (Garland, New York, 1995).
14. L. Stryer, *Biochemistry* (Freeman, New York, 1995).
15. A. Arkin, P. Shen, J. Ross, *Science* **277**, 1275 (1997); I. Schreiber, Y. F. Hung, J. Ross, *J. Phys. Chem.* **100**, 8556 (1996); A. Goldbeter, *Biochemical Oscillations and Cellular Rhythms: The Molecular Bases of Periodic and Chaotic Behavior* (Cambridge Univ. Press, Cambridge, 1996).
16. A. P. Alivisatos et al., *Adv. Mater.* **10**, 1297 (1998).
17. P. G. Wolynes, *Acc. Chem. Res.* **25**, 513 (1992).
18. W. Ehrfeld, L. J. Kricka, H. Lehr, *Top. Curr. Chem.* **194**, 233 (1998); G. Blankenstein and U. D. Larsen, *Biosens. Bioelectron.* **13**, 427 (1998).
19. S. T. Bramwell, P. C. W. Holdsworth, J. F. Pinton, *Nature* **396**, 552 (1998); V. S. L'vov, *ibid.*, p. 519.
20. P. Bak and M. Paczuski, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6689 (1995).
21. L. A. Thompson and J. A. Ellman, *Chem. Rev.* **96**, 555 (1996).
22. R. Beale and T. Jackson, *Neural Computing: An Introduction* (Hilger, New York, 1991); S. F. Zornetzer, J. L. Davis, C. Lau, Eds., *An Introduction to Neural and Electronic Networks* (Academic Press, San Diego, CA, 1990).
23. H. P. Lu and S. Xie, *Nature* **385**, 143 (1997); X. S. Xie and R. C. Dunn, *Science* **265**, 361 (1994); S. M. Nie and R. N. Zare, *Annu. Rev. Biophys. Biomol. Struct.* **26**, 567 (1997).
24. This work was supported by the Defense Advanced Research Projects Agency and by NIH grant GM30367. We acknowledge F. Frankel for providing Fig. 2.

VIEWPOINT

Complexity in Biological Signaling Systems

Gezhi Weng,¹ Upinder S. Bhalla,² Ravi Iyengar^{1*}

Biological signaling pathways interact with one another to form complex networks. Complexity arises from the large number of components, many with isoforms that have partially overlapping functions; from the connections among components; and from the spatial relationship between components. The origins of the complex behavior of signaling networks and analytical approaches to deal with the emergent complexity are discussed here.

Signaling in biological systems occurs at multiple levels. In its broad sense, one could use the term "signaling" to describe events ranging from interactions between single molecules to interactions between species in ecological systems. The aim here is to deal with complexity in signaling at a single level: intracellular signaling within a cell. We will outline how current

and forthcoming tools in biochemistry, cell and molecular biology, and physiology, as well as theoretical analysis and simulation methods, may be used to study this complex system.

In a general sense, the adjective "complex" describes a system or component that by design or function or both is difficult to understand and verify. In the past decade, analysis of complex systems (the field of complexity) has emerged as a distinct facet of mathematical and physical sciences. Understanding of biological systems may be enhanced by analysis of their complex nature. In physical systems, complexity is determined by such factors as the number of components and the intricacy of the interfaces between

them, the number and intricacy of conditional branches, the degree of nesting, and the types of data structures. Biological signaling networks possess many of these attributes, as well as dynamic assembly, translocation, degradation, and channeling of chemical reactions. All of these activities occur simultaneously, and each component participates in several different activities.

One approach to understanding complexity is to start with a conceptually simple view of signaling and add details that introduce new levels of complexity. As this process unfolds, it becomes clear where experimental data end and how progressively more difficult it becomes to understand the system as a whole in terms of the functional details of individual components.

A Signaling Wire

The simplest description of signaling may be in terms of elementary chemistry in a homogeneous well-stirred cell where all molecules have equal access to each other. Here, the most up-

¹Department of Pharmacology, Mount Sinai School of Medicine, New York, NY 10029, USA. ²National Center for Biological Sciences, UAS-GKVK Campus, Bangalore 560065, India.

*To whom correspondence should be addressed. E-mail: iyengar@mssm.edu