Development of Kinetic Models in the Nonlinear World of Molecular Cell Biology

Robert D. Phair

Increasingly, successful research on metabolic systems relies on teams of specialists. Because of the enormous complexity of these systems, many experimental groups have sought collaborations with theoreticians for data analysis and modeling. Predictably, cultural differences in scientific approach, methodology, assumptions, and language have led to some persistent difficulties in communication across the experiment-theory frontier. This report attempts to diagnose some of these difficulties from the perspective of 30 years' experience in both experimental and theoretical biology, and to suggest guidelines for effective collaboration between experimentalists and theorists. As these collaborations move to the level of cellular and molecular biology, effective communication will become all the more important because the simple linear rate laws of radiotracer and stable-isotope kinetics will no longer suffice. This is because every form of regulation and control, hallmarks of metabolic systems, results in nonlinear kinetics. To advance this transition to nonlinear cellular and molecular metabolic models and to facilitate communication between experimental and theoretical collaborators, a general procedure for incorporating control mechanisms in metabolic rate laws is developed based on the familiar rapid-equilibrium assumption of classical enzyme kinetics.

IN MANY AREAS OF BIOMEDICAL research, there is an uneasy alliance forming between theoretical and experimental biology. The alliance is uneasy for the same reasons encountered in other disciplines: the vocabulary and syntax are different, the values are different, and the experience is different. Yet these alliances are forming with greater regularity than ever before. One explanation for this trend is that theoretical biology is beginning to recognize that a theory is merely an untested hypothesis until its predictions are compared against experimental data, and experimental biology is faced daily with the need to comprehend an intertwined array of cellular regulatory and signaling systems whose complexity has far outstripped the capacity of the human mind.

Thus, the theorist and the experimentalist are drawn together. As in any alliance, effective communication is critical. Predictably, however, both professional groups come to joint discussions with their arcane dialects and jargon. Worse, they bring their own truths that are agreed upon within the group but are unstated and untested in mixed company. These seemingly inescapable human traits are significant barriers to effective communication in any alliance, and the theorist-experimentalist alliance is no exception.

These barriers block bilateral communication nearly as effectively as C.P. Snow's *Two Cultures* block communication between the sciences and the humanities. The result is described in Fig 1. This figure depicts an all-too-frequent paradigm in which the alliance forms with a single major goal: publication. Generally, this worthwhile goal is in fact reached, but the absence of effective communication within the collaborative alliance robs its members of nearly all of the other benefits that might be realized. Instead, the experimental investigators' deep knowledge of the data is lost, the unanswered questions of experimenters and methodologists are set aside, and a mere table of numbers is passed (often electronically) to the theoretician or modeler. In turn, the modeler treats these numbers as "truth" and proceeds to extract as many subtle and profound truths from them as is humanly possible. To do so, the theoretician adopts a number of assumptions, some explicit and some implicit, that make the problem tractable. These assumptions are rarely communicated to the experimental team. The overall result of this strangled communication is that only data and conclusions cross the chasm between bench investigation and theoretical analysis. A publication emerges from the effort, but it rarely has enough impact on the minds of the experimenters to modify their approach to the next experimental design. This is unfortunate. Without this feedback, the only benefit of collaboration is a publication whose conclusions have little bearing on the future course of experiment.

One approach to reducing these barriers between theorists and experimentalists is to reiterate the contents of that "black box," to make its walls transparent. Figure 2 is one conception of the approach taken by modelers and other theorists.[2] Because the diagram is closely related to classic scientific method, all scientists, whether experimentalists, theorists, or both, should be able to see themselves traveling along the arrows of Fig 2. Most of this diagram is familiar, but the role of the model may require some explanation. If you divide the diagram in half with a horizontal line, the top represents the world of theory and the bottom the world of experiment. Mathematical and computer models serve the theorist in nearly the same way that laboratory methods serve the experimentalist.

A computer model is an explicit realization of one hypothesis about how nature works. It permits the theorist to make quantitative predictions about how the living system will respond to a particular experimental protocol if this particular hypothesis is true. Similarly, the laboratory methods and protocols designed by the experimentalist permit him or her to ask nature a question and obtain a useful quantitative answer if the experimental design is effective.

Parallelism is also apparent in obtaining predictions from models and obtaining data from methods. In both cases, human intervention is intentionally minimal. Just as the experimentalist endeavors to prevent bias at this step, so the theorist adopts

From the Departments of Biomedical Engineering and Physiology, The Johns Hopkins University School of Medicine, Baltimore, MD

Submitted May 17, 1997; accepted May 19, 1997.

Presented in part at the National Institutes of Health Symposium, "In Vivo Tracer Kinetics and Modeling," May 3, 1996, National Institutes of Health, Bethesda, MD.

Address reprint requests to Robert D. Phair, PhD, BioInformatics Services, 12114 Gatewater Dr, Rockville, MD 20854.

0026-0495/97/4612-0021$03.00/0

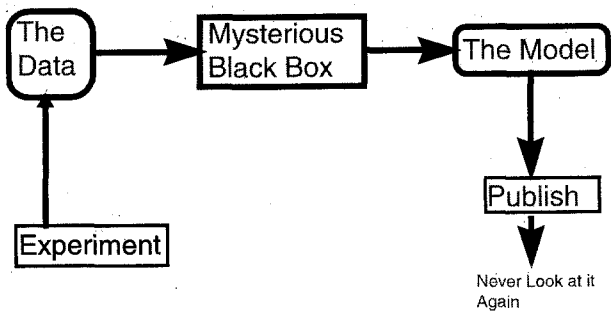

Fig 1. Ineffective collaboration between the theorist and the experimentalist. From the perspective of the experimentalist, it often appears that the data are being processed by a mysterious black box, presumably well understood by the theorist, whose output is "the model." When the theorist pronounces the model complete, a jointly authored report is produced, and unfortunately, the modeling effort is rarely called upon to inform the ongoing process of experimental design. In the worst case, the model has no impact at all on the direction of the experimental laboratory.

algorithms whose validity and lack of bias has been examined by all the methods available to modern numerical analysis. Interestingly, nonspecialists see these two processes, simulation and experiment, as the real work of modelers and experimentalists. The reality, however, is very different. The real work of the experimentalist is in experimental design and method development, and the real work of the theorist is in hypothesis creation and model development.

This is not to say that experimentalists do not create hypotheses and theorists do not design experiments—far from it. Indeed, the most effective collaborations involve both experimentalists and theorists in both hypothesis creation and experimental design. On the other hand, methods development is nearly always the domain of experimentalists and model development is nearly always the domain of theorists. Hence, it is model development and testing that holds most of the mystery

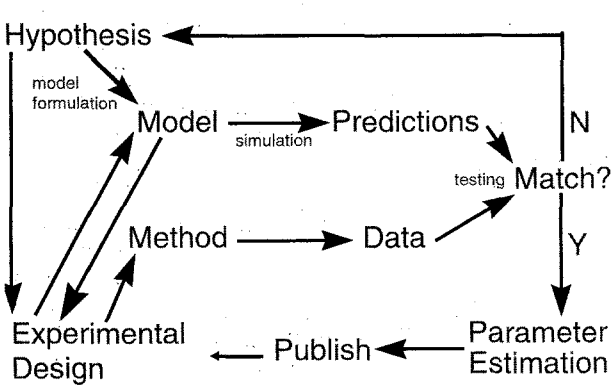

Fig 2. Inside the black box. A model is constructed as a precise statement of a working hypothesis or theory. This construction process is called "model formulation" and consists of transcribing the theory into a rigorous symbol and arrow diagram, and then translating the diagram into the language of mathematics using standard physical chemical principles. Almost always, the resulting system of equations cannot be solved by classic analytical (paper and pencil or symbolic math) methods. Consequently, the modeler adopts the tools of numerical analysis to obtain the model's prediction for a given experimental design. This numerical process is referred to as simulation.

in the mysterious black box of Fig 1, and it is the purpose of this report to dispel some of those mysteries in the context of cellular-level metabolic modeling.

A PRACTICAL PROCESS FOR MODEL DEVELOPMENT AND TESTING

Another way of reading Fig 2 is to interpret it as a graphical representation of the steps taken in model development and testing. One approach to this process, based on three decades' experience, can be summarized in six steps. First, assemble the experimental data. Second, record the experimental protocols in precise terms. Third, transcribe the theory or hypothesis into a rigorous system diagram. Fourth, translate the system diagram into the language of mathematics—construct the initial model. Fifth, impose the experimental protocols on the model using simulation. And sixth, compare the model predictions with the experimental data.

Typically, the initial model will fail to match the data. An anecdote will serve to illustrate this point. Between 1980 and 1996, I taught these methods to undergraduates and to graduate students in many of the basic medical sciences. These classes were project-based. Thus, in the course of 16 years, I supervised the construction of more than 200 models of biological systems. In each case, the model was constructed based on a published experimental report. Students formulated their initial models as quantitative statements of the researchers' published theory and then proceeded (as depicted in Fig 2) to compare the predictions of this theory with the published experimental data. Among these hundreds of models, there were only one or two for which parameter values could simply be adjusted to fit the data. All the rest of these projects required structural changes in the model (and therefore in the theory) to account for the data. The question thus becomes what to do upon encountering this mismatch.

When faced with discrepancies between the model and the data, modeling either earns its support or it disguises its discomfiture in several layers of impenetrable mathematics. In the face of discrepancies, the modeler must think as a biologist, not as a mathematician. She or he must make biological value judgments, and modify the model in some way that might account for the discrepancies. This distinction between thinking as a biologist and thinking as a mathematician is vital. I think it lies at the heart of all troubled relationships between theorists and experimentalists.

In some sense, it is always possible to add features (parameters and equations) to a model to account for any interesting feature of the experimental data. This follows from the notion of a series approximation to any function of time. For example, any time series of experimental observations measured in response to a physiological or pharmacological perturbation can be fitted to a sum of exponentials, and concordance between the data and the equation can always be improved by adding a new term to the series approximation. Nevertheless, few investigators would try to infer mechanistic significance from the coefficients and exponents of this series. Every feature of a mathematical or computer model of a biological system falls somewhere in this spectrum: it may represent a known physical chemical process taking place in a known subcellular compartment, it may represent a hypothesized chemical signal or

process, or it may have no connection, real or hypothesized, with physicochemical reality. In this last instance, the feature is entirely equivalent to the addition of another exponential term; it remains forever untestable because it has no link, not even a hypothesized one, to the measurable physical world.

When modeling severs the link to the measurable world, it discards fully half the power of modeling. Such models will still account beautifully for the experimental data, but they have abandoned the link to mechanism. Mathematical models lacking mechanistic links can be tremendously useful in diagnosis, but they can only guess about therapy. In my view, the most effective models are mechanistic in all details.

One additional illustration may help to cement this concept. Metabolism is not the only discipline that attracts modelers. Epidemiology, for example, has a vigorous literature filled with regression models, and the correlations discovered in these analyses are often assigned far-reaching causal interpretations. A famous example is the negative correlation between the plasma high-density lipoprotein cholesterol concentration and the frequency of a subsequent myocardial infarction. In concert with the positive correlation for low-density lipoprotein cholesterol, this mathematical model has been used to make thousands of diagnoses. Because it is not mechanistic, therapy has been targeted at plasma lipids as a surrogate for the disease process. This may work well, but is only a guess. It is a simple matter to imagine a "therapeutic" agent capable of producing a dramatic elevation in plasma high-density lipoprotein cholesterol that is nevertheless disastrous for the patient.

Returning now to the procedure for model development and testing, we see that we have made one circuit around the upper loop of Fig 2. The modeler must traverse this path as often as necessary to discover a model whose predictions display no gross systematic deviations from the available data. One can then ask how to give the hypothesis the best chance to account for the data. The current best answer to this question is to use an optimizer. The theory of optimization and parameter estimation is a vigorous and fascinating one with an extensive literature of its own,[3] and metabolic modeling relies on expert software development just as experimenters rely on expert engineers of specialized equipment such as ultracentrifuges, scintillation counters, and mass spectrometers.

Having completed this process, what you have is a precisely formulated hypothesis that is quantitatively consistent with your data. You also have estimates and measures of uncertainty for all of the parameters of your model. Of course, you will only have precise estimates when the data actually contain the required information.

Next, you apply the Zech Principle.[4] Use your model as a statement of current understanding and join forces with the experimentalist to design experiments to test it further. There are more detailed descriptions of the steps in the modeling process at http://www.webcom.com/rphair/bis/resources/cybertext/pkm-cont.html.

A natural history of kinetic modeling in any single area of metabolism begins with measurement of mass or concentration in two or more steady states. The next level of complexity is to measure tracer kinetics in steady states. Tracer methods derive their power from the ability to provide estimates of the basic kinetic parameters, rate constants (or clearances) and production rates, even though the experimental system is at steady state.[5,6] The next level of complexity is measurement of mass or concentration in non–steady states. This approach is commonplace in much of biomedical science and has spawned a large contingent of physiological modelers. In fact, a strong case can be made that understanding non–steady states is the ultimate goal of kinetic analysis, and experience suggests that new investigators would be wise to concentrate on this arena. The final frontier, not widely exploited as yet, consists of measuring tracer kinetics in non–steady states.[7,8] Metabolic experiments using stable isotopic tracers often fall into this category. Analysis of these data sets, with tracer dynamics superimposed on tracee dynamics, can be daunting, but new software tools are becoming available that have the potential to help extract useful fundamental information from extraordinarily complex data sets.

A central feature of models built to describe molecular cell biology is nonlinearity. A metabolic process is nonlinear when the flux through that process is not simply proportional to the concentration of a single metabolite. Instead, the flux depends on the concentrations of multiple metabolites. Thus, nonlinearity is easily seen to be the rule rather than the exception. The reason for this lies in a central biological concept—regulation. Scan the list of words in Table 1. The vocabulary of modern biological journals is rife with these and related words. Regulation and control are at the core of current biological and medical research, and adding any of the features in Table 1 to a computer model of a biological system leads quickly to a system of nonlinear differential equations.

As theorists and experimentalists begin to venture into the nonlinear world of molecular cell biology, they will need a practical procedure for including control mechanisms (such as those in Table 1) in their rate laws. The next section suggests such a procedure.

A PROCEDURE FOR INCLUDING CONTROL IN MECHANISTIC RATE LAWS

Experience suggests that the best starting point is construction of a rigorous diagram for the mechanism whose rate law you wish to construct. The modeler can then take advantage of the extensive and insightful literature on enzyme kinetics (for review, see Segel[9]) to construct useful mechanistic rate laws for a wide variety of protein-mediated processes. Moreover, it is usually sufficient to construct a rate law based on what is termed the rapid-equilibrium assumption. Briefly, this asserts that all of

Table 1. Features of Molecular Cell Biology That Require Nonlinearity in Computer Models

Regulation	Saturation
Control	Desensitization
Activation	Down(up)regulation
Inhibition	Adaptation
Allosterism	Inactivation
Binding	Initiation
Cooperativity	Modification
Synergism	Termination
Feedback	Repressors
Signaling	Transcription factors
Transduction	Induction

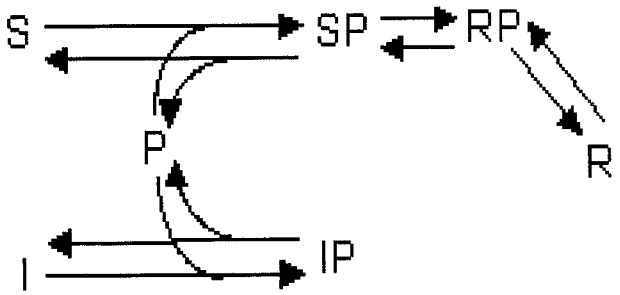

Fig 3. Symbol and arrow diagram for competitive inhibition. Substrate, S; product, R; protein, P; complexes, SP and RP; inhibitor, I. The catalytic or rate-limiting step is the transition from SP to RP. When the overall equilibrium constant does not overwhelmingly favor the formation of R, it is necessary to treat this step as reversible, as in this example. The 3 other reaction steps are treated as rapid equilibria; their detailed kinetics are treated as rapid compared with the SP to RP transition.

the regulators of the process whose rate law you are constructing are ligands that are in rapid equilibrium with their specific binding sites on the protein that mediates the process. Having drawn a diagram based on this principle, it becomes possible to transcribe the nonlinear rate law by inspection.

This process is best taught by example, and we begin with the familiar case of competitive inhibition. The first step is to transcribe the symbol and arrow diagram representing the mechanism you want to include in your model. The diagram for competitive inhibition is shown in Fig 3.

Under the rapid-equilibrium assumption, the rate of change of the product, dR/dt, is determined entirely by the net flux through the catalytic or transformation step, ie, the step in which SP is converted to RP. Among enzyme kineticists, equation 1 is termed the velocity equation:

$$\frac{dR}{dt} = k_{cat+}[SP] - k_{cat-}[RP]. \tag{1}$$

The next step uses the familiar mathematical device of multiplying the right side by a quantity equal to 1, which introduces some useful terms. The quantity is

$$\frac{P_{tot}}{[P] + [SP] + [RP] + [IP]}. \tag{2}$$

This ratio is equal to 1 because all of the protein exists in one of the four forms included in the denominator. Next, we apply the definitions of forward and reverse maximal velocities:

$$\begin{aligned} V_{max+} &= P_{tot}k_{cat+}, \text{ and} \\ V_{max-} &= P_{tot}k_{cat-}. \end{aligned} \tag{3}$$

Substituting these into equation 1 yields the rate of appearance of product in terms of the two maximal velocities:

$$\frac{dR}{dt} = \frac{V_{max+}[SP] - V_{max-}[RP]}{[P] + [SP] + [RP] + [IP]}. \tag{4}$$

Now, we make use of the rapid-equilibrium assumption for all three of the ligands, S, R, and I. We use equilibrium binding relations to express every concentration in terms of P and the corresponding equilibrium constant:

$$\frac{dR}{dt} = \frac{V_{max+}\frac{[S][P]}{K_S} - V_{max-}\frac{[R][P]}{K_R}}{[P] + \frac{[S][P]}{K_S} + \frac{[R][P]}{K_R} + \frac{[I][P]}{K_I}}. \tag{5}$$

Finally, cancel [P] from the numerator and denominator, and obtain the classic rate law for competitive inhibition:

$$\frac{dR}{dt} = \frac{V_{max+}\frac{[S]}{K_S} - V_{max-}\frac{[R]}{K_R}}{1 + \frac{[S]}{K_S} + \frac{[R]}{K_R} + \frac{[I]}{K_I}}. \tag{6}$$

This method is useful in a wide variety of metabolic situations. It can be used to formulate nonlinear rate laws for all of the processes listed in Table 1. As a second example, we consider the omnipresent biological process, activation.

For activation, we proceed as before. First, transcribe the symbol and arrow diagram as shown in Fig 4. This activator mechanism works by modifying the substrate to make it a more effective substrate for the catalytic protein. The activator, A, binds to the substrate, S, so that the protein, P, now has two candidate substrates, the unmodified S and the activator-substrate complex AS. As before, we next write the velocity equation,

$$\frac{dR}{dt} = k_+[SP] + \beta[k_+[ASP]], \tag{7}$$

where β is the factor by which ASP is more rapidly converted to product than SP. If we apply the same steps used for the case of inhibition, we obtain a rate law useful for modeling the effect of an activator,

$$\frac{dR}{dt} = \frac{V_{max}\frac{[S]}{K_S} + \beta\left[V_{max}\frac{[A][S]}{K_{AS}K_D}\right]}{1 + \frac{[S]}{K_S} + \frac{[A][S]}{K_{AS}K_D}}, \tag{8}$$

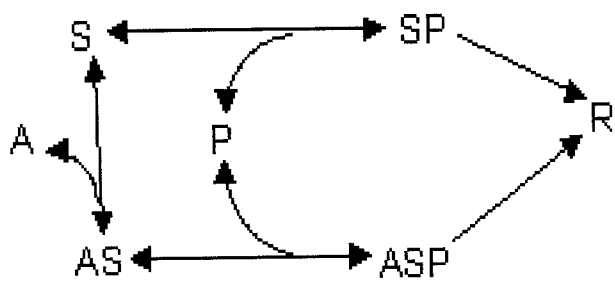

Fig 4. Symbol and arrow diagram for a process controlled by an activator. Substrate, S; activator, A; activator-substrate complex, AS; protein, P; protein complexes, SP and ASP; product, R. This derivation considers the case of an overall equilibrium constant that strongly favors formation of R. The Haldane relationship then ensures (so long as the substrate constants are comparable) that the reverse process will have a negligible flux compared with the forward process. Consequently, arrows in the diagram are unidirectional for production of R. Note also that an activator provides an alternative metabolic route to R.

where K_S is the equilibrium constant for the binding of S to P, K_D is the equilibrium constant characterizing the dissociation of the AS complex, and K_{AS} is the equilibrium constant for AS binding to P.

An Example From Cellular Endocrinology

Having seen how rate laws for activation and inhibition can be constructed, it may be useful to apply the method to a specific cellular process. For this, we can imagine ourselves in the position of a cellular biologist studying the role of the Bip chaperone protein in the control of peptide hormone secretion. The investigator's current question is how to construct a rate law for translation of new hormone and delivery to the lumen of the endoplasmic reticulum (ER). This would permit him or her to examine the kinetics of hormone translocation in a cell line that expresses a mutant sec63, which binds Bip poorly.

The working hypothesis we want to test is depicted in Fig 5. Among many ER proteins and macromolecular complexes involved in this process, several are central to the investigator's working hypothesis as shown in the figure. This mechanism is clearly more complex than the simple inhibition and activation examples of Figs 3 and 4, but the rapid-equilibrium assumption still applies. We start by writing the rate law based on the two "catalytic" steps represented by bold arrows in Fig 5:

$$\frac{dH}{dt} = k_{cat1}[mRNs61s63Bip] + k_{cat2}[mRNBips61s63]. \quad (9)$$

In equation 9, k_{cat1} and k_{cat2} are the catalytic rate constants for the process of translocating newly synthesized peptide hormone to the ER lumen. To reduce the length of the equations, the names of complexes and proteins have been shortened. The derivative, dH/dt, on the left side represents the rate at which new hormone is delivered to the ER lumen. In a complete model, there would also be a negative term on the right side of equation 9 representing vesicular transport of the hormone from the ER to cis Golgi, but we are just interested in constructing a term representing the flux of newly translated hormone.

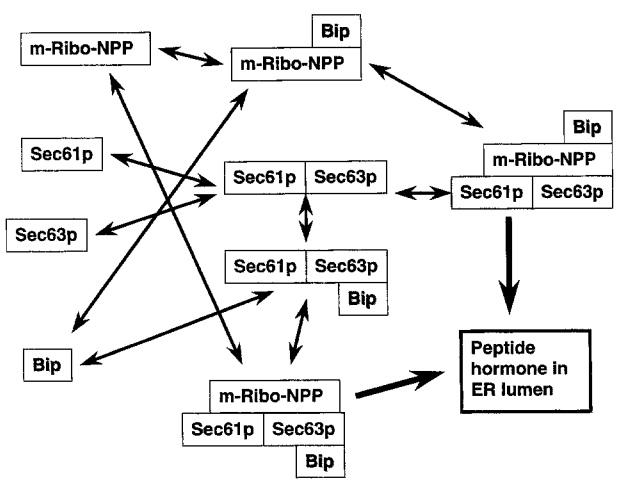

Fig 5. Symbol and arrow diagram of a hypothetical mechanism of peptide hormone translocation into the rough ER and its regulation by the chaperone protein, Bip. Sec61p and sec63p are proteins thought to form the aqueous pore in the ER membrane that permits transit of the nascent polypeptide (NPP). The complex of mRNA, ribosome, and nascent polypeptide is labeled m-Ribo-NPP. Double-headed arrows represent binding reactions treated as equilibria under the rapid-equilibrium assumption. Two bold unidirectional arrows represent translocation of the completed peptide hormone into the ER lumen; these are the "catalytic" steps.

Next, proceeding as before, we multiply the right side of equation 9 by the ratio,

$$\frac{sec61tot}{s61 + s61s63 + s61s63Bip + mRNBips61s63 + mRNs61s63Bip}. \quad (10)$$

Sec61 is chosen because it is the protein "catalyzing" the translocation of protein from the ribosome to the ER lumen. Applying the definitions of maximal velocity, the multiplication yields,

$$\frac{dH}{dt} = \frac{V_{max1}[mRNs61s63bip] + V_{max2}[mRNBips61s63]}{[s61] + [s61s63] + [s61s63Bip] + [mRNBips61s63] + [mRNs61s63Bip]}. \quad (11)$$

Now, we again apply the rapid-equilibrium assumption to express the concentration of each molecular complex in terms of the concentrations of its components and the appropriate equilibrium constants:

$$\frac{dH}{dt} = \frac{V_{max1}\dfrac{[s61][s63][Bip][mRN]}{K_D K_{B63} K_{R61}} + V_{max2}\dfrac{[s61][s63][Bip][mRN]}{K_D K_{R61} K_{BN}}}{[s61] + \dfrac{[s61][s63]}{K_D} + \dfrac{[s61][s63][Bip]}{K_D K_{B63}} + \dfrac{[s61][s63][Bip][mRN]}{K_D K_{R61} K_{BN}} + \dfrac{[s61][s63][Bip][mRN]}{K_D K_{B63} K_{R61}}}. \quad (12)$$

In equation 12, K_D is the equilibrium constant for sec63 binding to sec61, K_{B63} is the equilibrium constant for binding of Bip to sec63, K_{R61} is the equilibrium constant for binding of the 60S ribosomal subunit to sec61, and K_{BN} is the equilibrium constant for Bip binding to the nascent polypeptide hormone. Canceling [s61], we arrive at the desired rate law:

$$\frac{dH}{dt} = \frac{V_{max1}\dfrac{[s63][Bip][mRN]}{K_D K_{B63} K_{R61}} + V_{max2}\dfrac{[s63][Bip][mRN]}{K_D K_{R61} K_{BN}}}{1 + \dfrac{[s63]}{K_D} + \dfrac{[s63][Bip]}{K_D K_{B63}} + \dfrac{[s63][Bip][mRN]}{K_D K_{R61} K_{BN}} + \dfrac{[s63][Bip][mRN]}{K_D K_{B63} K_{R61}}}. \quad (13)$$

We now have a nonlinear rate law for the flux of new hormone being translocated into the ER lumen. The cell line expressing a mutant sec63 will require a larger value for the constant, K_{B63}, corresponding to a reduced affinity. This will be detected as a decreased rate of appearance of new hormone in the ER fraction.

This method is readily generalized to processes mediated by any protein and regulated by any number of activating or inhibiting ligands, some of which may also be proteins (as in the example). Because the method presented here is based in biochemistry and thermodynamics, it may serve as a natural language shared by theorists and experimentalists as they begin joint explorations of cellular and molecular metabolic systems.

COLLABORATION IN MODELING

The software tools for metabolic modeling are evolving quickly.[10,11] They are becoming easier to use and much more approachable for biologists with little or no advanced mathematical training. Consequently, it is entirely feasible that more and more experimentalists will decide to include kinetic modeling in their personal array of scientific tools.

This possibility raises a key question. What factors should a biologist consider in deciding whether to commit time and resources to becoming an expert modeler? One way to approach this question is to list the advantages and disadvantages as seen from the perspective of experience.

There are significant advantages to constructing the model in your own laboratory. First, this approach permits optimal communication of biological value judgments and theoretical limitations. Second, a new experiment suggested by today's modeling work can often be set up at the bench tomorrow. This daily feedback between model and experimental design is invaluable. Third, an investigator who has mastered both model formulation and experimental design develops a keener understanding of the power of modeling. He or she is well positioned to identify novel hybrid approaches to important questions. Fourth, there are no difficulties with intellectual property; the entire effort is clearly associated with a single laboratory.

The disadvantages of in-house modeling can also be enumerated. First, like all skills, modeling takes time. Since time is the most inelastic of resources, modeling will also require dilution of the investigator's effort in some other realm. Second, additional training or workshop experience is essential to apply modeling effectively. Third, new methods always have attendant hardware and software problems, and modeling is no exception. Fourth, both clinical investigation and experimental biology are difficult endeavors as it is; modeling can add significant value to the final product, but it is not simple.

GUIDELINES FOR COLLABORATION AND PARTNERING

Our economy continues to move toward specialization and outsourcing; someone who specializes in a given area can almost always do a job faster, at lower cost, and to a higher standard of quality than a nonspecialist. In light of this and having considered the disadvantages outlined in the previous section, a biomedical investigator might very reasonably choose to gain the advantages of modeling through collaboration or partnering with an expert.

Experience is an excellent, if expensive, teacher, and the following guidelines may thus be helpful to both theorists and experimentalists contemplating a cooperative effort.

Principle 1

Start with the experimentalist's theory. If only data change hands, the collaboration is weakened. Instead, the experimentalist should provide his or her working hypothesis at the same time he or she provides the data. The theorist should insist on this, too, and the working hypothesis should be the first one tested. This is true not because the working hypothesis is necessarily the simplest one that accounts for the data, but because it is bound to be a compact description of an enormous body of knowledge and experience that the theorist cannot afford to ignore. There may well be portions of this theory that cannot be resolved from the current data set, but this does not mean they should be eliminated from the model. A reasonable alternative, which allows the collaboration to build on what has been learned in the past, is to fix some parts of the model based on data from other experiments or other laboratories.

Principle 2

Detail the details. Both collaborators should expect to ask and answer a great many detailed questions about experimental protocols and methods. If this conversation does not take place, the collaboration is flawed. Reasons for attending to detail are many, but perhaps the most important is the theorist's need to model the experiment as well as the biological system. As Fig 2 illustrates, comparison of model predictions to experimental data depends on performing the same experiment on the model as was performed on the living system. This requirement enjoins the collaborators to define carefully what was done and when so that, in effect, both are performing the same experiment.

Principle 3

Educate your collaborators. Ensure that the modeler understands the relevant biomedical jargon and the experimentalist understands the relevant theoretical jargon. Theorists who do not comprehend recovery calculations or quenching or antibody specificity and experimentalists who do not comprehend the significance of a rate constant or the difference between steady state and equilibrium are taking unnecessary professional risks. They risk flatly incorrect conclusions, and they risk the inability to present and defend substantial portions of their joint work.

Principle 4

Agree on a time for project completion. This is a basic principle of project management. In an effective collaboration, the experimentalist should be prepared to deliver all of the data and detailed descriptions of the experimental protocols in a timely fashion. He or she should also arrange to answer questions quickly and completely as collaborative work proceeds. In return, the theorist should be prepared to commit to a schedule for completion of data analysis. Every research project is under pressure to be productive, and theorists cannot afford to claim an exemption from this economic reality. Everyone knows that unforeseen events cause deadlines to slip, but this is not justification for abandoning deadlines.

Principle 5

Anticipate intellectual property issues. At the start of a team project, it should be clear how each team member will be rewarded and what each will contribute to earn that reward. The wise team member will always contribute far more than expected, but prudent investigators make basic expectations

explicit at the start. Specify the author list or lists for the reports that will be written and who will be expected to write which sections. Discuss what future publication rights all authors have with respect to the model and the data used to test it. This discussion may be difficult to initiate, but the resulting collaboration will be more businesslike, productive, and sustainable.

Principle 6

Expect modeling to test, not necessarily to support, your favorite hypothesis. In an effective collaboration, experimentalists are prepared to suggest alternative hypotheses when their favorite theory is shown to be inconsistent with their own data. Both theorists and experimentalists should be party to hypothesis creation. Without this constant internal feedback, the experimentalist risks having to defend an untenable position, and the theorist risks postulating a new mechanism that is inconsistent with published results from multiple laboratories worldwide.

Principle 7

Insist on mechanism. A minimalist definition of a mathematical model is a system of equations that fit the available data. Effective collaborations are built on mechanistic models, and not all mathematical models are mechanistic. By insisting that each and every feature of the model correspond to a known or hypothesized biological or physical chemical mechanism, the vital link between theory and experiment is maintained. If the model contains elements that do not correspond to measurable quantities, it becomes an untestable, and therefore unscientific, hypothesis. It may fit the data beautifully, but without the link to mechanism, one may as well have opted for a simple transfer function or a sum of exponentials.

These seven principles are based on 30 years' experience in both experimental and theoretical work. They will surely need modification in individual circumstances, but they can serve as a good foundation for effective collaboration and partnering between experimental and theoretical biologists.

REFERENCES

1. Snow CP: The Two Cultures. Cambridge University Press, 1993

2. Phair RD: The role of kinetic modeling in physiology, in Berman M, Grundy SM, Howard BV (eds): Lipoprotein Kinetics and Modeling. New York, NY, Academic, 1982

3. D'Argenio DZ, Schumitzky A: A program package for simulation and parameter estimation in pharmacokinetic systems. Comput Programs Biomed 9:115-134, 1979

4. Fisher WR, Venkatakrishnan V, Fisher ES, et al: The ^{3}H-leucine tracer: Its use in kinetic studies of plasma lipoproteins. Metabolism 46:333-342, 1997

5. Zech LA, Boston RC, Foster DM: The methodology of compartmental modeling as applied to the investigation of lipoprotein metabolism. Methods Enzymol 129:366-384, 1986

6. Barrett PHR, Foster DM: Design and analysis of lipid tracer kinetic studies. Curr Opin Lipidol 7:143-148, 1996

7. Caumo A, Homan M, Katz H, et al: Glucose turnover in presence of changing glucose concentrations: Error analysis for glucose disappearance. Am J Physiol 269:E557-E567, 1995

8. Caumo A, Zerman A, Rizza R, et al: The dual tracer time-varying volume method for measuring hepatic glucose release in nonsteady state: Theoretical and simulation results. Comput Methods Programs Biomed 41:243-267, 1994

9. Segel I: Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems. New York, NY, Wiley, 1975

10. Zech LA, Rader DJ, Greif PC: Berman's simulation analysis and modeling. Adv Exp Med Biol 285:189-199, 1991

11. Foster DM, Barrett PHR, Beltz WF, et al: Using SAAM II to Model Kinetic and Pharmacokinetic Data. Proceedings of the Simulation in Health Sciences Conference. San Diego, CA, Society for Computer Simulation, 1994, pp 87-90