0026-895X/96/010086-06$3.00/0

MOLECULAR PHARMACOLOGY, 50:86–95 (1996).

KINFIT II: A Nonlinear Least-Squares Program for Analysis of Kinetic Binding Data

G. ENRICO ROVATI, RICHARD SHRAGER, SIMONETTA NICOSIA, and PETER J. MUNSON

Laboratory of Molecular Pharmacology, Institute of Pharmacological Sciences, University of Milan, 20133 Milan, Italy (G.E.R., S.N.), and Laboratory of Applied Studies (R.S.) and Analytical Biostatistics Section (P.J.M.), Division of Computer Research and Technology, National Institutes of Health, Bethesda, Maryland 20892

Received February 28, 1996; Accepted March 29, 1996

SUMMARY

We describe a versatile computer program for least-squares fitting of ligand/receptor association and dissociation curves from several experiments simultaneously. The program is designed to handle any number of classes of binding sites reacting with a single ligand that may have two forms: labeled and unlabeled. For a single class of binding sites, the exact, analytical solution is used to generate the computed curves. For more than one class of sites, the computed curves are generated through numerical solution of a set of ordinary differential equations. The parameters determined with this procedure are the on- and off-rate constants and the concentrations of each binding site class. An extensive selection of experimental designs can be processed. The times of observation may be freely chosen, although that choice will affect the quality of the results. Starting with a sample material (e.g., cells, membranes, macromolecules), one can preincubate (to equilibrium) with any combination of labeled and unlabeled ligand. One can then perturb the system by adding any combination of labeled and cold ligand or simply diluting the sample; such an experiment can be continued through several perturbations. A variety of such runs may be combined for analysis as a single data set.

We present a new computer program, KINFIT II, for the simultaneous analysis of several kinetic ligand-binding experiments. The program estimates association rates (on rates), dissociation rates (off rates), and binding capacities or concentrations of one or more classes of binding sites in a sample. KINFIT II is written in MATLAB 4, an interactive language for scientific numerical calculations (1).

Ligand-binding studies are still one of the most widely used techniques in many laboratories, from pharmacology to endocrinology and toxicology or biochemistry, for characterizing the interaction of ligands (e.g., hormones, drugs, neurotransmitters) with a wide variety of receptors, macromolecules, membranes, or whole cells. Binding studies are typically divided into two main categories: equilibrium and kinetic studies. Here, we primarily consider kinetic studies.

Basically, there are two types of kinetic experiments possible: the association and the dissociation time course. In an association experiment, labeled ligand is added to a solution containing receptor, and aliquots of the mixture are taken at chosen times. In a dissociation experiment, the sample is usually preincubated to equilibrium with labeled ligand. Then, unlabeled ligand or simply a diluting solution is added, and aliquots are taken. The association and dissociation experiments can be done in tandem; i.e., first add the labeled ligand, take aliquots over time, and then dilute the same sample and take more aliquots. The tandem design provides more information than either experiment alone while eliminating the need for a separate preincubation step.

In general, a run (a single time course after a given perturbation) can continue an earlier one (i.e., a tandem design) by perturbing a system that might or might not yet be at equilibrium. The perturbation itself can consist of any combination of added labeled and unlabeled ligand or simply a dilution.

Kinetic ligand binding experiments are very often used in preliminary phases of the characterization of a receptor system. Data from dissociation time courses are also used to calculate the dissociation constant(s) (i.e., k_{off}) and to test whether the data from the dissociation kinetic experiment suggest heterogeneity of binding sites.

In the simplest case of a single class of binding site, eqs. 1 and 2 (see Methods) simplify to a Ricatti differential equation that has an exact analytical solution (2). It is possible to solve these equations for H_1, the bound ligand concentration, with k_{on}, k_{off}, and R_1 as unknown parameters (3). When two or more binding sites are present, no analytical solution is available except in the special case of a large excess of H_T

This work was partially supported by the Italian National Research Council, Committee for Information Science and Technology.

ABBREVIATIONS: ODE, ordinary differential equation; LT, leukotriene.

and/or C_T, so we solve it numerically. This avoids the need to use a pseudo-first-order approximation (4), which may be of questionable validity.

Although it is theoretically possible to estimate all three parameters (k_{on}, k_{off}, and R) from a single dissociation experiment, there may be practical advantages to analyzing association and dissociation curves together. In particular, greater confidence can be asserted in the correctness of the model when it fits data taken from both experimental protocols.

Methods

The following quantities are fixed by nature: n is the number of site classes; k_{on_i} is the on-rate constant for the ith site class, $i = 1, \ldots, n$; k_{off_i} is the off-rate constant for the ith site class; $K_{d_i} = k_{off_i}/k_{on_i}$, the equilibrium binding constant for ith site class; and S_i is the concentration of ith binding site class, in original solution, in user-defined units.

The next quantities may be manipulated experimentally but are fixed during each phase of the experiment: c is the scale factor, multiplying S_i; ϕ is the dilution factor incurred at the start of each phase due to the addition of solution containing additional reagents (see Appendix B); $R_i = cS_i$, the total molar concentration of the ith receptor or binding site class; H_T is the total molar concentration of labeled ligand; and C_T is the total molar concentration of unlabeled ligand.

The following quantities vary during the time course of an experiment: $H_i(t)$, the molar concentration of labeled ligand bound to class i sites at time t; $C_i(t)$, the molar concentration of unlabeled ligand bound to class i sites at time t; $H_F(t)$, the molar concentration of free labeled ligand at time t; and $C_F(t)$, the molar concentration of free unlabeled ligand at time t.

We use the following additional notation: df (degrees of freedom), WSS (weighted sum of squares), CV (coefficient of variation), δ (parameter dependency), and t (time).

In the absence of unlabeled ligand, the system is governed by n ODEs of the following form:

$$dH_i(t)/dt = k_{on_i} H_F(t)[R_i - H_i(t)] - k_{off_i} H_i(t), \quad i = 1,2,\ldots,n \qquad (1)$$

where

$$H_F(t) = H_T - \sum_{j=1}^{n} H_j(t)$$

When the system contains mixtures of labeled and unlabeled ligand, eq. 1 must be modified and augmented with n more ODEs, yielding the following:

$$\begin{aligned} dH_i(t)/dt &= k_{on_i} H_F(t)[R_i - C_i(t) - H_i(t)] - k_{off_i} H_i(t) \\ dC_i(t)/dt &= k_{on_i} C_F(t)[R_i - C_i(t) - H_i(t)] - k_{off_i} C_i(t), \quad i = 1,2,\ldots,n \end{aligned} \qquad (2)$$

where

$$C_F(t) = C_T - \sum_{j=1}^{n} C_j(t).$$

When $n = 1$, the single ODE (for the labeled-only case) or the two ODEs (for the labeled-and-unlabeled case) can be solved analytically. The solution was given by Rodbard and Weiss (2), but there have been some errata (see Appendix A). When $n > 1$, eqs. 1 and 2 must be solved numerically. We use Gear's second- and third-order formulas to integrate a stiff system of ODEs (5).

These equations model both association and dissociation experiments. For association experiments, we specify the initial condition $H_i(0) = 0$ (i.e., that all labeled ligand is initially unbound.) Dissociation experiments require that labeled ligand be preincubated with receptor and that this preequilibrated condition be perturbed by dilution with or without the addition of unlabeled ligand to the system. The combination of preincubation and perturbation conditions determine the values of H_T, C_T, and the initial conditions, $H_i(0)$ and $C_i(0)$, for eq. 2. Here, we model experiments combining an association phase with a dissociation phase, initiated at a certain time by the dilution and addition of unlabeled ligand. KINFIT II will be used to analyze other, more complex protocols, as described in Appendix B. An association/dissociation experiment is described with the use of four numbers: H_T for the association phase ($C_T = 0$ during association), a dilution factor ϕ (which modifies both H_T and R_i for the dissociation phase), and C_T for the dissociation phase and the time at which dissociation is initiated.

The data from both association and dissociation experiments may be fitted simultaneously with the use of KINFIT II. Likewise, curves obtained with different labeled or unlabeled ligand concentrations (H_T and C_T) may also be fitted simultaneously. The mathematical framework and numerical methods for fitting functions nonlinear in the parameters to experimental data have been discussed in detail previously (6). The method used here is a weighted nonlinear least-squares curve-fitting routine with the use of the Gauss-Newton algorithm as modified by Levenberg and Marquardt (7, 8). The partial derivatives of the fitting function with respect to the unknown parameters (e.g., k_{on_i}, k_{off_i}, and R_i), required by the Gauss-Newton technique, are estimated numerically.

To simulate an experimental situation, we must specify the characteristics of the system under study (n, k_{on_i}, k_{off_i}, and R_i for each i). We must also specify the experimentally controllable parameters (H_T and C_T) in each phase of the experiment and the time at which the next phase is begun. Furthermore, we must specify the times at which observations of H_i are made. With eq. 1 or 2, the "true" values of the experimental observations can be simulated. Finally, we assumed a gaussian distribution for measurement error with a variance proportional to the measured quantity (constant CV). In our study, we assumed a constant 1% error rate.

After producing simulated data, we fit them with eq. 1 or 2 as appropriate with the correct number of binding classes (n) with the use of KINFIT II. This exercise obviously should lead to parameter values close to those used in the simulation step. Any departures from ideal values reflect the limitations imposed by the experimental design (e.g., number and spacing of time points, choice of ligand concentrations).

The quality of the estimated parameters can be conveniently summarized with the use of two quantities that are available after any fit to real or simulated data. The first quantity is the standard error of each parameter, a measure of the spread or uncertainty of the value. We express this as a proportion of the parameter value estimate (i.e., the CV). The CV is directly proportional to the magnitude of simulated experimental errors; thus, a reduction (simulated or real) in experimental error will reduce the CV of all estimated parameters. Useful estimates should have a CV of around $\leq$30%.

The second quantity is δ, which is used to summarize the interactions or correlations of a parameter estimate with all other parameters in the model. The squared δ (δ^2) is also known as the *variance inflation factor* in the statistical literature (9). Both measure how much higher the parameter error becomes due to the presence of other unknown and correlated parameters in the model. Values for δ of >100 suggest that there are serious problems with either the model or the experimental design (see Appendix B). Thus, experimental designs (e.g., choice of measurement time points, H_T and C_T values for each phase) should be sought that minimize δ and the parameter standard errors.

In any model, there are implicit physical assumptions that should be recognized. In the model described with the use of eqs. 1 and 2, these assumptions are that (a) labeled and unlabeled ligands have the same rate constants. This assumption holds well in systems in

which labeled and unlabeled ligands differ only in the isotopes they contain. The introduction of a radioactive iodine into a protein, however, may modify the constants considerably. (b) There is no cooperativity within or between site classes. This assumption is not appropriate for, e.g., hemoglobin/oxygen kinetics. (c) Ligand does not self-associate or bind with anything (e.g., buffers) except the sites in question. (d) Likewise, the sites do not bind anything but ligand. (e) There are no transient effects of binding that would alter the rate constants or concentrations (e.g., the Bohr effect). (f) The system is well stirred. When aliquots are taken, the volume is reduced without the concentration changing. (g) The bound ligand or bound fraction can be accurately determined at known time points in the reaction.

These assumptions are often met sufficiently well for modeling purposes, whereas otherwise they must serve as a useful first approximation.

When additional parameters are added to a model (i.e., an increase in the number of binding classes), the fit usually improves, resulting in a smaller WSS simply because of the added flexibility in the model. However, the resulting model might be unnecessarily complex with an excess of parameters and loss of *df*. Therefore, a compromise is necessary between the addition of new parameters and improvement in the goodness of fit. The appropriate number of parameters (i.e., the appropriate number of binding sites), can be determined with a statistical test, the F test, based on the extra sum of squares principle (10) that compares the more complex model (involving more parameters) with the simpler one.

Results

We provide some examples that demonstrate the capabilities of the program and that partially explore the limits and possibilities of kinetic binding experiments. To do this, we used synthetic data obtained from a given model of known complexity and with known parameter values. We considered cases with either one or two classes of binding sites with a single labeled ligand concentration for the association phase (cases 1 and 2) and a case with two binding sites and an experiment consisting of two association/dissociation curve pairs run at different labeled ligand concentrations (case 3). We set the number of points and time schedule for each curve in the simulations so that both the association and dissociation curves almost reach equilibrium. For convenience, we express the binding affinity (K_d) in molar concentration units and the off rate (k_{off}) in per-minute units. The binding capacity of the system will always be expressed as a unitless ratio to the binding affinity (R/K_d), and, likewise, the total ligand concentration will be given in terms of K_d.

We also analyzed two real kinetic experiments with the use of KINFIT II. In the first experiment, [^{3}H]ketanserin binding to human platelets is modeled as a single homogeneous class. Ketanserin is a potent serotonin antagonist, acting at the type 2 receptor (5-hydroxytryptamine$_2$), but we demonstrated (11) in human platelets that this compound actually binds to the monoamine transporter rather than to the 5-hydroxytryptamine$_2$ receptor. Analysis of both equilibrium and dissociation binding data clearly demonstrated the presence of a homogeneous class of sites. In the second experiment, [^{3}H]LTC$_4$ binding to U937 cells is analyzed with the use of a two-site model. Leukotrienes are potent inflammatory mediators, and their receptors and metabolism have been extensively studied in the cell line U937, an immortalized line of human promonocytic leukemia cells. In particular, we found through the use of a equilibrium binding assay that in these cells, [^{3}H]LTC$_4$ binds not only to a well-characterized low affinity/high capacity site, which is most likely an enzyme (12), but also to a high affinity/low capacity site with the characteristic of a specific receptor.[1] Therefore, the kinetic experiment was run twice (in the presence of different labeled and unlabeled ligand concentrations), a design that should facilitate the estimation of all six parameters of the model.

TABLE 1

Influence of the labeled total ligand concentration, H_T on the precision of parameter estimates from an association/dissociation experiment (case 1)

The simulated parameters were k_{off} = 1 min^{-1}, R = 0.1 K_d, C_T = 1 mM for the dissociation phase, and dilution factor ϕ = 0.9.

k_{on}	K_d	H_T	Dependency δ		
			k_{on}	k_{off}	R
nM^{-1}	nM	K_d			
10	0.1	0.01	>100	5.9	>100
10	0.1	0.1	90	5.6	94
10	0.1	1.0	8	3.6	10
10	0.1	10	1.4	1.4	1.8
10	0.1	100	1	1	1.1
1	1.0	0.01	>100	5.9	>100
1	1.0	0.1	90	5.6	94
1	1.0	1.0	8	3.6	10
1	1.0	10	1.4	1.4	1.8
1	1.0	100	1	1	1.1
0.1	10.0	0.01	>100	6.4	>100
0.1	10.0	0.1	87	5.9	91
0.1	10.0	1.0	7.7	3.6	10
0.1	10.0	10	1.4	1.4	1.8
0.1	10.0	100	1	1	1.1

Case 1: Single association/dissociation curve, one class of sites. In the association phase of this simulation, all of the labeled ligand is initially unbound [$H_i(0) = 0$, $C_i(0) = 0$]. The dissociation phase is initiated by dilution of the system by a modest factor (ϕ is 9:10 = 0.9) together with the addition of unlabeled ligand (C_T = 1 mM). The simplest model includes a single class of binding sites. Therefore, three parameters must be considered: k_{on}, k_{off}, and R. We consider a series of parameter settings (Table 1). For each setting, we ran a series of simulations while varying the total ligand concentration, H_T, from 0.01 to 100 K_d for the association phase. For the dissociation phase, C_T became 1 mM, whereas H_T and R_i were reduced by the addition of a solution containing unlabeled ligand. We simulated 40 data points for each phase (association and dissociation) of the experiment that were spread uniformly over a time interval, which allowed the system to essentially reach equilibrium. It is evident from Table 1 that as the H_T concentration becomes smaller than that of K_d, δ rises for k_{on} and R but not substantially for k_{off}.

In a second set of simulations of one association and one dissociation curve, we varied the R values with a constant K_d. Table 2 shows the influence of binding capacity, expressed as R/K_d, on the precision of the parameter estimates. When the binding capacity is >10 K_d, the δ values for k_{on} and R begin to rise, becoming essentially infinite when R reaches 1000 K_d. Again, as expected, there is no effect on the dependency of the dissociation constant.

We also considered a situation in which instead of using a large excess of unlabeled ligand to start the dissociation phase, we added only a small amount of the unlabeled ligand.

[1] G. E. Rovati and S. Nicosia, unpublished observations.

TABLE 2

Influence of the binding capacity, R, on the dependency of parameter estimates from an association/dissociation experiment (case 1)

The model parameters were set as K_d = 1 nM and k_{off} = 1 min^{-1} (k_{on} = 1 nM^{-1} min^{-1}). Labeled ligand concentration H_T = 10 K_d for the association phase, C_T = 1 mM for dissociation phase, and dilution factor ϕ = 0.9.

Binding capacity R	Dependency δ		
	k_{on}	k_{off}	R
K_d			
0.01	1.4	1.4	1.8
0.1	1.4	1.4	1.8
1	1.5	1.4	1.9
10	2.5	1.7	3.2
100	37	1	37
1000	>100	1	>100

In this case, the dissociation was incomplete (Fig. 1), but our computer program, taking into account the concentrations of both the labeled and unlabeled ligand, can easily handle this experimental protocol. As can be seen from CVs and δ values (see legend to Fig. 1), we can easily calculate all of the unknown parameters of the model even if the dissociation phase is incomplete, thus using only a fraction of the amount of unlabeled ligand that is usually added to completely dissociate the system.

In Fig. 2A, we present a real set of data obtained with the use of [^{3}H]ketanserin in human platelet membranes. In this real case, we used 3.5 nM labeled ketanserin, and we ran an association experiment, taking seven aliquots at different times. In a second, separate experiment, we again incubated membranes from human platelets with 3.5 nM [^{3}H]ketanserin and let the system equilibrate for 15 min before adding 158 μM unlabeled ketanserin to completely dissociate the system and taking aliquots. Simultaneous computer analysis

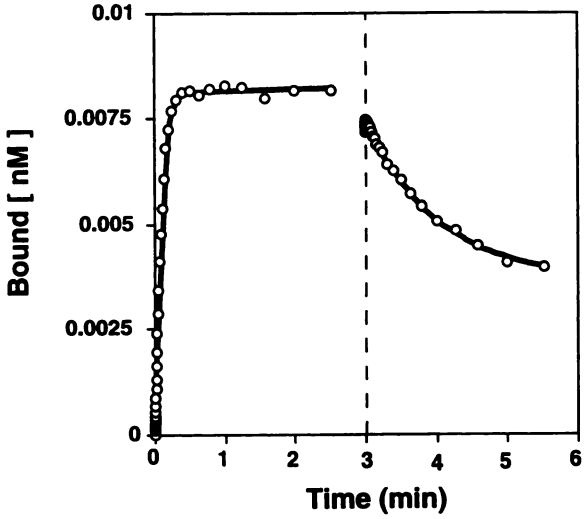

Fig. 1. Association and dissociation curves and data simulated by a kinetic model of a single class of binding sites. *Vertical dashed line*, the addition of unlabeled ligand. The model parameters were set as K_D = 0.1 nM, k_{off} = 1 min^{-1} (k_{on} = 10 nM^{-1} min^{-1}), and R = 0.01 nM. At 3 min, the dissociation phase was initiated by the dilution and addition of unlabeled ligand. H_T (association phase) = 10 K_d, C_T (dissociation phase) = 1.0 nM, and dilution factor ϕ = 0.9. The parameter estimates differed from their true values by ≤2%. The dependencies, δ, of the estimated parameters were 1.5, 2.2, and 2.7, whereas the estimated CV values were 1, 2.2, and <1 for k_{on}, k_{off}, and R, respectively.

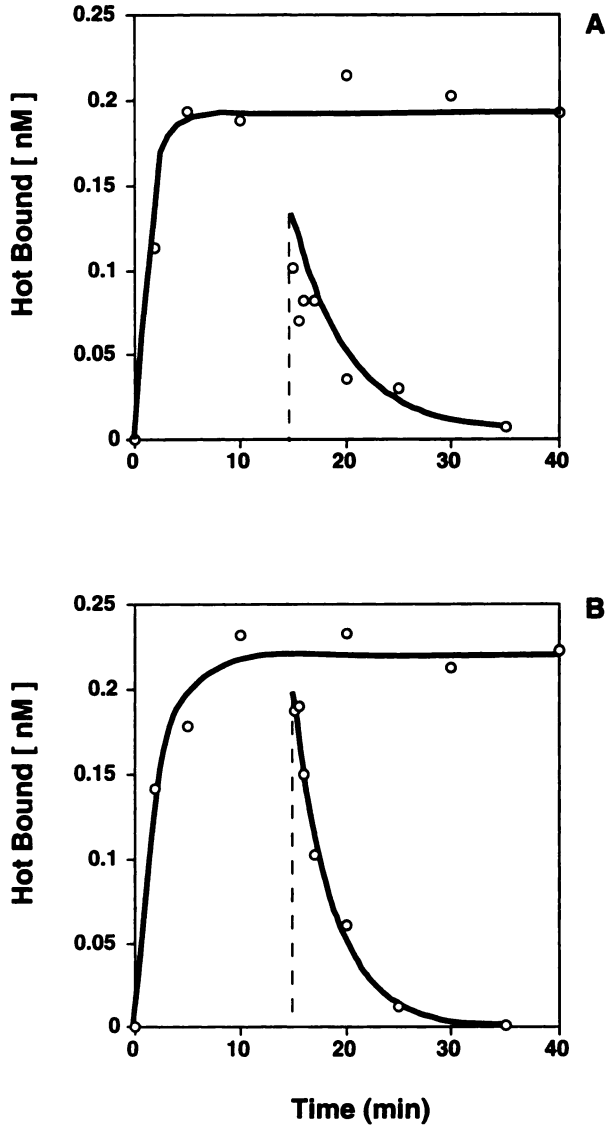

Fig. 2. Association curve and dissociation curve of ketanserin. A, Real experimental data. The labeled ligand concentration H_T used for the association curve was 3.5 nM [^{3}H]ketanserin, whereas the dissociation phase was initiated at 15 min (*dashed line*) by the addition of unlabeled ketanserin (dissociation phase C_T = 158 μM). Parameter values estimated from the real data are k_{on} = 0.054 nM^{-1} min^{-1} (±86% CV), k_{off} = 0.26 min^{-1} (±31% CV) (K_d = 4.9 nM), and R = 0.62 nM (±65% CV). B, Synthetic data generated with 10% random error in labeled bound on the basis of the same time schedule and parameter as A. Parameter values estimated from the synthetic data are k_{on} = 0.051 nM^{-1} min^{-1} (±42% CV), k_{off} = 0.27 min^{-1} (±11% CV) (K_d = 5.3 nM), and R = 0.63 nM (±31% CV).

of both of these sets of data with the use of KINFIT II gives a k_{on} estimate of 5.36 × 10^7 M^{-1} min^{-1} (CV = 86%, δ = 13.3), a k_{off} value of 0.26 min^{-1} (CV = 31%, δ = 4.8), and an R value of 6.2 × 10^{-10} M (CV = 66%, δ = 16.7). The corresponding estimate for K_d was 4.9 nM, which is practically identical to the K_d estimated separately with the use of an equilibrium binding study (4.79 nM, CV = 45%) (11).

In Fig. 2B, we also present a set of synthetic data generated with 10% random error in labeled bound on the basis of the same time schedule and parameter estimates obtained from the analysis of the real data presented in Fig. 2A. It is

clear that real and synthetic data are in good agreement, thus confirming the reliability of our mathematical model.

Case 2: Single association/dissociation curve, two classes of sites. In this set of simulations, we considered the situation in which two classes of binding sites are present. Therefore, six parameters should be estimated from data: k_{on_1}, k_{off_1}, and R_1 for the first binding site and k_{on_2}, k_{off_2}, and R_2 for the second. We considered cases in which two sites are present in equal amounts or in which the low affinity site is present in a higher amount.

When the two-site classes are present in equal amounts, we set a 10-fold difference in the K_d values and simulated models with five different H_T concentrations in the association phase. Based on the information in Table 3, it is clear that the class of binding sites with higher affinity for the ligand is better defined than the one with lower affinity.

In Table 4, we considered a series of models with different capacity ratios (R_2/R_1) for the two sites, with the K_d values set at 10 R for both sites and H_T values set at 10 K_{d_2}. Not surprisingly, as the ratio increases, the estimates of the lower capacity binding site quickly become indeterminate.

A special situation may be considered when two classes of binding sites have the identical K_d values but different k_{off} values and, thus, different values for $k_{on} = (k_{off}/K_d)$. We ran such a simulation and analyzed the results by using KINFIT II. The program was able to precisely identify all six different parameters with a single experiment, even though different proportions of binding sites were assumed (Fig. 3).

Case 3: Multiple association and dissociation curves, two classes of sites. To solve the problem of parameter indeterminacy discussed in case 2, we suggest the use of multiple association and dissociation curves run at several initial ligand concentrations, H_T. To illustrate this idea, we simulated two association and dissociation experiments run at a different ligand concentration. We choose an H_T concentration for each experiment to be 10-fold higher than the K_d values for each of the two sites. We varied the binding capacity R_2 from 1 × R_1 to 1000 × R_1. Now, we were able to estimate all six parameters with reasonable precision, provided R_2 remains at <100 R_1, and in all cases the precision was greater than when only a single association/dissociation experiment was run (compare models 1 and 2 of Table 4 with models 2 and 3 of Table 5). This demonstrates the potential benefit of the more extensive experimental protocol. In other investigations, we considered similar protocols for systems with a variety of affinity ratios (K_{d_2}/K_{d_1}) and binding capacities.

When the K_{d_1} differs from K_{d_2}, three different possibilities arise: (a) both k_{on} and k_{off} of the two sites are different, (b) only k_{on_1} and k_{on_2} are different, or (c) only k_{off_1} and k_{off_2} are different. We considered the possibility that different K_d values arise from differences in the association rates (Fig. 4) or from differences in dissociation rates (Fig. 5). For this set of simulations, we fixed the binding capacity ratio R_2/R_1 at 10, with a constant R/K_d ratio of 0.1 for each class of sites. The simulated experiments shown in Figs. 4 and 5 consist of two association/dissociation pairs, with the association-phase total ligand concentration $H_T = 10 \times K_{d_1}$ for the first experiment and $H_T = 10 \times K_{d_2}$ for the second. With the program, we were able to precisely determine all six parameters for these experiments, with CV values of <10% and δ values of <15. Such precise estimates would not have been possible with only one pair of association/dissociation curves (one H_T concentration).

In other simulations, we examined the behavior of the

TABLE 3

Effects of total labeled ligand concentration, H_T, on precision of estimates when model has two classes of binding sites with equal capacity. Experiment involves a single pair of association/dissociation curves (case 2)

The simulated parameters were set as K_{d_1} = 0.1 nM, k_{off_1} = 1 min^{-1} (k_{on_1} = 10 nM^{-1} min^{-1}), and R_1 = 0.1 K_{d_1} and K_{d_2} = 1 nM, k_{off_2} = 5 min^{-1} (k_{on_2} = 5 nM^{-1} min^{-1}), and $R_2 = R_1$.

Model	H_T	Dependency δ					
		k_{on_1}	k_{off_1}	R_1	k_{on_2}	k_{off_2}	R_2
	K_{d_1}						
1	0.1	>100	15	>100	>100	17	>100
2	1	14	16	62	>100	29	>100
3	10	9.4	4.7	56	>100	12	>100
4	100	2.7	3.6	30	4.2	4.8	23
5	1000	9.9	3.2	27	10	3.4	25

TABLE 4

Effect of binding capacity, R_2, for the second class of sites on precision of estimates (case 2)

The simulated parameters were set as K_{d_1} = 0.1 nM, k_{off_1} = 1 min^{-1} (k_{on_1} = 10 nM^{-1} min^{-1}), and R_1 = 0.01 nM and K_{d_2} = 10 R_2 and k_{off_2} = 5 min^{-1}.

Model	H_T	R_2	CV					
			k_{on_1}	k_{off_1}	R_1	k_{on_2}	k_{off_2}	R_2
	K_{d_1}	R_1				%		
1	100	10	17	25	29	2.5	3	2.8
2	1000	100	>100	>100	70	1.8	1.3	0.9
3	10000	1000	>100	>100	80	1	0.9	0.2

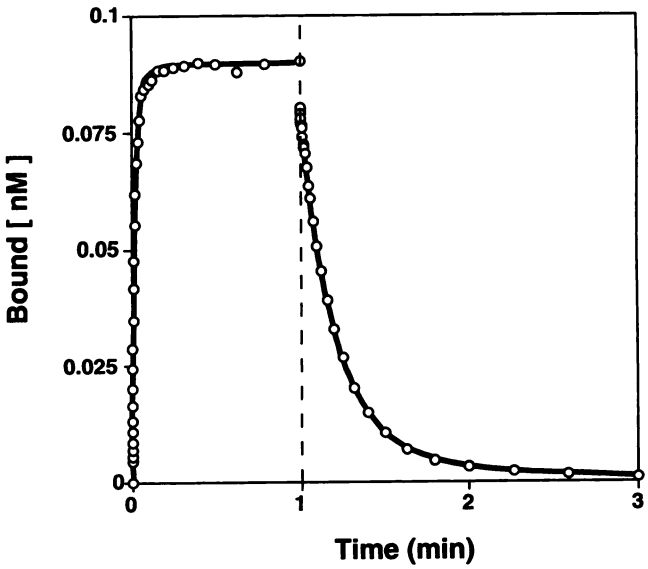

Fig. 3. Association and dissociation curves simulating a model with two different classes of binding sites with the same equilibrium dissociation constants ($K_{d_1} = K_{d_2}$ = 0.1 nM) but different kinetic constants. *Vertical dashed line*, the addition of unlabeled ligand. The model parameters were k_{on_1} = 10 nM^{-1} min^{-1}, k_{off_1} = 1 min^{-1}, R_1 = 0.01 nM, k_{on_2} = 50 nM^{-1} min^{-1}, k_{off_2} = 5 min^{-1}, R_2= 0.1 nM, H_T = 1 nM, C_T = 1 mM, and ϕ = 0.9. The estimated values of the model were k_{on_1} = 11.4 nM^{-1} min^{-1}, k_{off_1} = 1.2 min^{-1} (K_{d_1} = 0.105 nM), and R_1 = 0.013 nM and k_{on_2} = 51 nM^{-1} min^{-1}, k_{off_2} = 5.1 min^{-1} (K_{d_2} = 0.1), and R_2 = 0.097 nM, with CV values of 4.1%, 3.8%, and 2.2% and of 3.9%, 3.8%, and 2.3%, respectively.

TABLE 5

Influence of the binding capacity, R_2, on CV for experiment including two sets of association/dissociation curves at different H_T values (case 3)

The simulated parameters were set as K_{d_1} = 0.1 nM, k_{off_1} = 1 min^{-1} (k_{on_1} = 10 nM^{-1} min^{-1}), R_1 = 0.01 nM, and k_{off_2} = 5 min^{-1} and H_T (first experiment, association phase) = 1 nM = 10 K_{d_1} and H_T and second experiment, association phase) = 10 K_{d_2}.

Model	Binding capacity R_2	Binding affinity K_{d_2}	CV					
			k_{on_1}	k_{off_1}	R_1	k_{on_2}	k_{off_2}	R_2
	R_1				%			
1	1	1	2.3	3.3	3.7	2.5	3.9	3.5
2	10	1	11	16	18	1.2	1.9	1.6
3	100	10	70	57	26	0.9	0.7	0.2
4	1000	100	>100	>100	>100	0.7	0.7	0.2

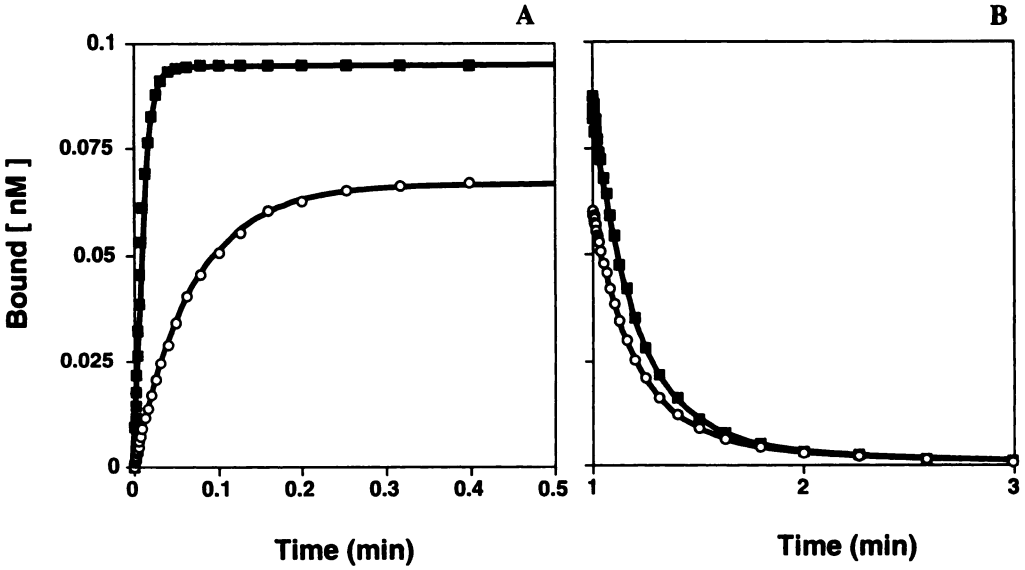

Fig. 4. Kinetic experiments run at different total ligand concentrations. In the association phase, H_T = 1 nM (○) and H_T = 10 nM (■); in the dissociation phase, C_T = 1 mM and ϕ = 0.9 for both curves. Association (A) and dissociation (B) curves simulate a model with two different classes of binding sites having the same dissociation constants (k_{off_1} = k_{off_2} = 1 min^{-1}) but different association constants. For the simulation, parameters were k_{on_1} = 10 nM^{-1} min^{-1}, R_1 = 0.01 nM, k_{on_2} = 1 nM^{-1} min^{-1}, and R_2 = 0.1 nM.

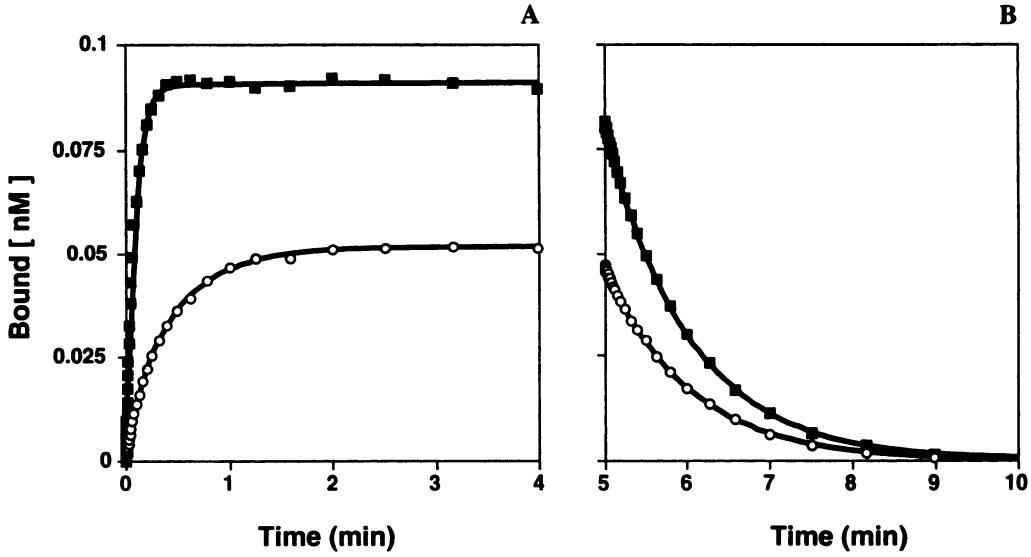

Fig. 5. Kinetic experiments run at different total ligand concentrations. In the association phase, H_T = 1 nM (○) and H_T = 10 nM (■); in the dissociation phase, C_T = 1 mM and ϕ = 0.9 for both curves. Association (A) and dissociation (B) curves simulate a model with two different classes of binding sites having the same association constants (k_{on_1} = k_{on_2} = 10 nM^{-1} min^{-1}) but different dissociation constants. The simulated parameters were k_{off_1} = 1 min^{-1}, R_1 = 0.01 nM, k_{off_2} = 10 min^{-1}, and R_2 = 0.1 nM.

estimation strategy with the use of a variety of binding capacity ratios. In all eight data sets tested with the program, we were able to determine the six unknown parameters of the simulated models, with CV values of <63%.

In Fig. 6, we present two different experiments obtained with [^{3}H]LTC_4 in membranes from U937. In this real case, we used two association and two dissociation curves of [^{3}H]LTC_4 run at different total ligand concentrations (0.5 nM [^{3}H]LTC_4 for the first experiment and 1 nM [^{3}H]LTC_4 plus 30 nM LTC_4 for the second experiment). Both dissociation phases were initiated by 3 μM unlabeled LTC_4. Simultaneous computer analysis of both of these sets of data with the use of KINFIT II according to one- and two-site models gave a statistically significant improvement in the two-site model compared with the simpler one ($p < 0.01$). The calculated parameters were as follows: k_{on_1} = 0.24 nM^{-1} min^{-1} (CV =

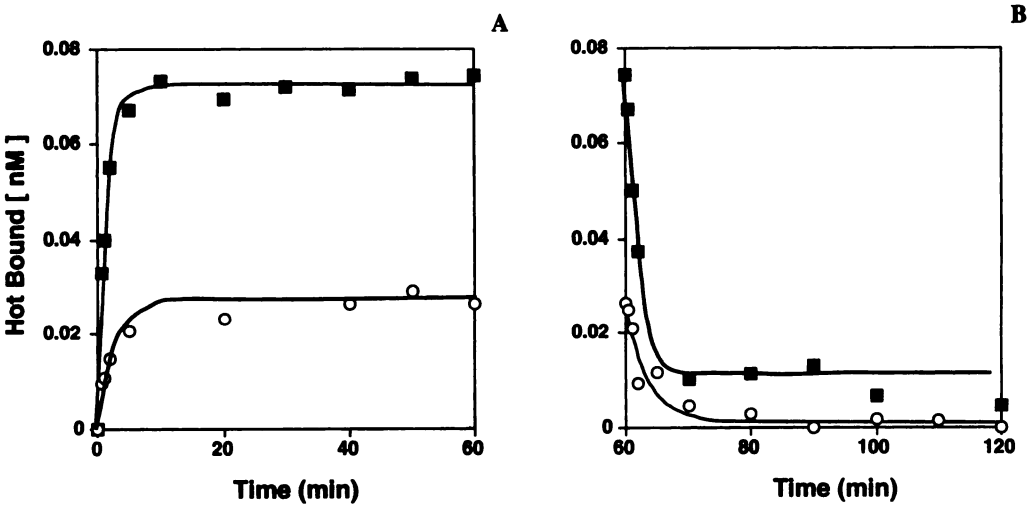

Fig. 6. Two association (A) and two dissociation (B) curves of [^{3}H]LTC$_4$ run at different total ligand concentrations: ○, association phase H_T = 0.5 nM [^{3}H]LTC$_4$; ■, association phase H_T = 1 nM [^{3}H]LTC$_4$; C_T = 30 nM LTC$_4$. Both dissociation phases were initiated by C_T = 3 μM unlabeled LTC$_4$. Parameter values estimated simultaneously from the four curves were k_{on_1} = 0.24 nM^{-1} min^{-1} (±34% CV), k_{off_1} = 0.25 min^{-1} (±19% CV), and R_1 = 0.069 nM (±28% CV) (K_{d_1} = 1.1 nM) and k_{on_2} = 0.0038 nM^{-1} min^{-1} (±19% CV), k_{off_2} = 0.51 min^{-1} (±8.5% CV), and R_2 = 7.5 nM (±16% CV) (K_{d_2} = 134 nM).

34%), k_{off_1} = 0.25 min^{-1} (CV = 19%), and R_1 = 0.069 nM (CV = 28%) (K_{d_1} = 1.1 nM) and k_{on_2} = 0.0038 nM^{-1} min^{-1} (CV = 19%), k_{off_2} = 0.51 min^{-1} (CV = 8.5%), and R_2 = 7.5 nM (CV = 16%) (K_{d_2} = 134.8 nM).

Discussion

We described a computer program for the analysis and modeling of kinetic ligand-binding experiments. Our algorithm is suitable for analysis of binding experiments involving one or more classes of binding sites and one or more curves, performed at different ligand concentrations. We presented some synthetic data sets simulating very common situations encountered in the real world as well as real data as examples of the power and flexibility of KINFIT II.

KINFIT II allows the estimation of parameters for a standard association experiment and a classic dissociation experiment, in which the dissociation is induced by either a very large dilution or the addition of a large excess of unlabeled ligand. In addition, KINFIT II can be used to extract the unknown parameters when only a small amount of unlabeled ligand is added to start the dissociation phase (Fig. 1). This feature gives the experimentalist the possibility of trying different experimental designs for the dissociation phase of a kinetic experiment and can result in a considerable saving of unlabeled ligand, which can be costly or limited in quantity.

We demonstrated that data from association and dissociation time courses can be used to model a receptor system; to estimate parameter values such as k_{on}, k_{off}, and R; or to statistically test alternative hypotheses regarding the structure of the model. Because KINFIT II can handle classic kinetic protocols (single association and dissociation curves) and because families of kinetic curves run at different ligand concentrations, it can effectively combine the power of kinetic and equilibrium studies to probe the nature of a complex receptor system.

We previously demonstrated that a homologous displacement curve or a saturation curve should start with a total ligand concentration as low as possible, consistent with obtaining reliable measurements (13). In other words, in equilibrium binding experiments, one has to start with a concentration of the labeled ligand as low as possible in both the displacement and saturation curves, whereas in kinetic experiments, the best precision and lowest dependency values are obtained with high concentrations of labeled ligand, H_T. This observation can be very useful when the labeled ligand has a very high affinity for a particular receptor class but low specific activity, as is often the case with tritiated substances. The use of very low ligand concentrations obviously creates problems when radioactive count rates drop into the background levels. We suggest that in this situation it might be advantageous to determine the equilibrium constant through a kinetic experiment, with sufficient labeled ligand to ensure adequate count rates for a reliable measurement (Table 1).

The test cases illustrate that when one of the two classes of binding sites greatly outnumbers the other, parameters associated with the minority class are less well resolved. Also, if the amount of ligand bound in one class greatly exceeds that bound in the other class, for the range of ligand concentrations used the same effect on resolution will be observed.

It is interesting to note that KINFIT II can resolve two sites with equal K_d values but different kinetic parameters with a single association and dissociation time course (Fig. 3). This is in contrast to the situation with equilibrium experiments, which cannot distinguish two such classes. This is one of the situations in which kinetic experiments have a superior resolution power compared with equilibrium binding experiments.

The most rigorous approach to the problem of characterizing multiple binding sites is, of course, to run more than one curve at different ligand concentrations and then analyze them simultaneously, as presented in case 3. Such combined experiments should be considered when validating a hypothetical receptor mechanism or when interest centers on both the kinetic and equilibrium characteristics of the system under investigation.

The test cases also show that with a wide variety of rates and site populations, the program was able to recover the "true" parameter values with a minimal degree of bias and acceptable variance. Two exceptions are noted. First, with two different classes of binding sites having a >100-fold difference in binding capacity, the use of a single H_T concentration is not sufficient to precisely estimate parameters associated with the lower capacity class (Table 4). Second, in the presence of two different classes of binding sites, even with the use of more than one ligand concentration (Table 5),

if the capacity of one class greatly outweighs the other, parameters associated with the lower-capacity class are not well resolved.

We also present two sets of real data obtained in a simple one-binding-site model and in a more complex, two-binding-site model. In the first case, we analyzed the data to obtain the model and the parameter values (Fig. 2A), and we double-checked them against the data obtained from equilibrium binding experiments, thus confirming the very good agreement between them. Furthermore, we used KINFIT II, given the parameter values and the time schedule of the real experiment, to back-calculate the synthetic data (Fig. 2B) that should have been obtained if the model and the calculated parameters were correct. It is clear from Fig. 2 that there is good correspondence between real and synthetic data, supporting our mathematical model and the interpretation of the results.

The second set of real data came from a system with two distinct binding sites for LTC_4 in membranes from U937 cells. As mentioned before, we previously validated the model with a series of equilibrium experiments. However, in this type of experiment, despite a clear indication of the existence of two distinct binding sites, the absolute value of the K_d of the high affinity/low capacity site was not precisely determined (K_d = 0.03–0.2 nM) due to the very high affinity but low specific activity (see above) of the labeled ligand. Therefore, we decided to perform kinetic experiments that should have a superior resolution power in this situation. From the simultaneous analysis of two kinetic experiments run at different total ligand concentrations, we were able to obtain a statistically significant improvement in the two-site model compared with the one-site model ($p < 0.01$), thus confirming and validating the model obtained with the equilibrium experiments. However, the kinetic experiments yielded an estimate of the K_d of the high affinity/low capacity site of ~1 order of magnitude different from the one obtained from equilibrium experiments (K_d = 1.1 nM). Thus, we believe that the >100-fold difference in capacity (R = 0.069 versus 7.5 nM for the first and second site, respectively) confirms the intrinsic limit of the kinetic experiments when dealing with two sites, one of which greatly outnumbers the other. On the other hand, the K_d of the second site was very similar in equilibrium and kinetic experiments (134 and 201 nM, respectively).

In conclusion, with the use of our program KINFIT II, it is possible to analyze kinetic data using the exact mathematical model for one or more classes of sites, thus taking advantage of all of the information contained in the association and dissociation time courses.

Appendix A

The following are errata to the analytical solution of eqs. 1 and 2 (as published in Ref. 2).

In eq. B6, within U^+, the expression $a^2 + b^2$ should be $a^2 - b^2$.

In eq. B13, the fraction $p/(1 - \alpha)$ should be $K_p/(1 - \alpha)$.

In eq. B14, the numerator $p(q - U\_)$ should be $K_p(q - U\_)$.

The solution was republished in altered and corrected form (14).

Appendix B

KINFIT II Input

KINFIT II accepts kinetic data and produces estimates for the three parameters k_{on}, k_{off}, and S_i for each site class $i = 1, \ldots, n$. The input data are *pin*, a column vector of initial parameter estimates; t, a column vector of times for one or more runs; h, a column vector of observed bound labeled ligand; w, a column vector of statistical weights; and *info*, a seven-column matrix, with one row for each run.

The vector *pin* serves two purposes. First, it contains initial estimates of the fitting parameters. Second, its length divided by 3 determines the number, n, of site classes to be resolved. For example, the MATLAB statement

$$pin = [\ 1, 3, 5, 4, 9, 2\ ]';$$

means that two-site classes are to be used, with the initial estimates of $k_{on_1} = 1$, $k_{off_1} = 3$, and $S_1 = 5$ and $k_{on_2} = 4$, $k_{off_2} = 9$, and $S_2 = 2$.

We note that the parameter R_i in eqs. 1 and 2 is identified with the product cS_i, where c is a scale factor (described below). These parameters are explained more thoroughly in KINFIT II output, below. The vector t likewise serves two functions. First, it contains the observation times relative to the most recent perturbation. Second, any $t_j = 0$ announces the beginning of a new run. For example, the MATLAB statement

$$t = [\ 0,1,2,4,\ 0,2,4,5,6,\ 0,1,2,3,5,10\ ]';$$

describes a data set with three runs of four, five, and six observations, respectively.

The vector h contains the concentration of labeled ligand bound to sites of all classes. This is the observed variable, h_j, corresponding to t_j. The elements of the vector w should be reciprocal variances of each h_j. Where h_j is measured by counting statistics, w_j should be proportional to reciprocal counts. Thus, each jth observation consists of the three numbers t_j, h_j, and w_j.

The matrix *info* defines amounts initially present and amounts added for each run. Every run specified by some $t_j = 0$ must have a corresponding row of *info*. Within each row of *info*, the elements are 1) V^o, initial volume, containing the following three quantities; 2) c, a correction or scale factor for binding site concentration; 3) H^o, initial labeled ligand concentration; 4) C^o, initial unlabeled ligand concentration; 5) V^A, added volume, containing the following two concentrations; 6) H^A, added labeled ligand concentration; and 7) C^A, added unlabeled ligand concentration; there are a total of seven columns. The final volume, V_T, and dilution factor, ϕ, for the run are then computed within the program as:

$$V_T = V^o + V^A,\ \phi = \frac{V^o}{V_T} \qquad (3)$$

The factor c may be entered in absolute (e.g., molar) or relative (dimensionless) units. The choice of units for c will influence the units in which the fitted concentrations S are expressed. This flexibility is required for cases in which the molecules with binding sites are embedded in membranes or cells where the mass of the sample is known but the concentration of binding molecules is not. The best one can do is to assume that the binding sites are well distributed through-

out the sample material. Implications for the fitted concentrations are discussed in KINFIT II output.

The condition $c > 0$ not only provides the scale factor for the concentration of sample but also signals that this is a new run; i.e., this run is not tandem to the previous run. In this case, the values R_i, H_T, and C_T in eq. 1 or 2 are set to cS_i, H^o, and C^o, respectively, and the preequilibrated bound ligand concentrations H_i^∞ and H_i^∞ are found so that all derivatives are zero. Then, all concentrations are scaled by the dilution factor $\phi = V^o/V_T$, incurred by the additional volume, so that in eq. 1 or 2, we set $H_T = \phi H^o + (1 - \phi)H^A$, $C_T = \phi C^o + (1 - \phi)C^A$, and $R_i = \phi cS_i$. The initial conditions are set to $H_i(0) = \phi H_i^\infty$ and $C_i(0) = \phi H_i^\infty$.

In contrast, $c = 0$ signifies that this run is tandem to the previous run. The values for eq. 1 or 2 are $R_i = \phi cS_i$, using the c from the previous run, and $H_T = \phi H^1 + (1 - \phi)H^A$ and $C_T = \phi C^1 + (1 - \phi)C^A$, where H^1 and C^1 are H_T and C_T from the previous run, respectively. The initial conditions $H_i(0)$ and $C_i(0)$ are taken from the corresponding values at the final time point of the previous run, scaled by the dilution factor ϕ. Several successive tandem runs may be indicated in this way.

The following matrix *info* will serve as an example:

```
                 %  Vº   c   Hº   Cº   Vᴬ   Hᴬ   Cᴬ
                 [  1,  8,  0,  0,  1,  8,  0;     %  Run  1
info        =       1,  0,  0,  0,  1,  0,  4;     %  Run  2
                    1,  0,  0,  0,  1,  0,  0;     %  Run  3
                    3,  2,  2,  0,  3,  0,  2  ];  %  Run  4
```

(MATLAB comments begin with the symbol % and are ignored by the interpreter.) Run 1 begins with no ligand present, so the initial conditions are $H_i(0) = 0$ for all i. Then, labeled ligand is added. At the onset of run 1 (after the dilution by $\phi = 0.5$), the total concentrations are $H_T = 4$ and $C_T = 0$. The total volume at the onset of run 1 is $V_T = 2$, but (presumably because aliquots are taken) by the start of run 2, only one unit is left. Run 2 is tandem to run 1, and unlabeled ligand is added. Run 3 is tandem to run 2, and the perturbation here is a simple dilution (added volume, no added ligand). By the onset of run 3, the total concentrations have been diluted to $H_T = 1$ and $C_T = 1$. Run 4 is a new run not related to run 3. Sample and labeled ligands are preequilibrated; then, unlabeled ligand is added. At the onset of run 4, the total concentrations are the same as in run 3. However, the kinetics are quite different in run 4 because $C_i(0) = 0$ and free unlabeled ligand is being introduced, whereas in run 3, free and bound unlabeled ligands are carried over from run 2.

KINFIT II Output

There are three output vectors listed beside each other in columns. These are *p*, the vector of parameters; *se*, the vector of corresponding standard errors; and *dep*, the vector of corresponding dependencies (measures of condition). The vector *p* contains the final estimates of the parameters k_{on}, k_{off}, and *S* and so on for as many sites as were specified in the input vector *pin*. The parameters k_{off} are the first-order off-rate constants in units of 1/time and will of course depend on the units in which time is expressed. However, the units of k_{on}, the second-order on-rate constants, are 1/[(concentration)(time)] and will therefore depend on both the concentration units and the time units that are used.

The products *cS* are given in concentration units. Therefore, if *c* (specified in the matrix *info*) is already in concentration units, the values for *S* are nondimensional. They can be interpreted as sites/molecule. However, if *c* represents relative concentrations of binding material, then the values for *S* are concentrations of sites/unit amount of binding material.

The vector *se* contains approximate estimates of the standard errors of *p*, as described previously (15). The qualification "asymptotic" must be used because a linear error analysis is being applied to a nonlinear fitting problem. The dependency values *dep* are likewise asymptotic values. Parameter dependencies serve to separate two aspects of the standard errors. Standard errors are approximately proportional to the noise level in the data. However, this certainly is not the only influence. The number of fitted parameters in the model and the range of data (i.e., the experimental design) also determine the standard errors, sometimes compensating for each other. For example, extension of the model to include more parameters often entails extending the data range to improve resolution.

Dependency values are a critique of how well we have done in this regard. Let σ_i be the ith standard error produced by the program, let δ_i be the ith dependency value, and let ϵ_i be the standard error that the program would have produced if p_i were the only fitted parameter, with all other fitted parameters regarded as constants. Like σ_i, ϵ_i is proportional to noise level and serves as a measure of how well these data resolve this parameter in isolation. The dependency value is a ratio of the two standard errors: $\delta_i = \sigma_i/\epsilon_i$ or $\sigma_i = \delta_i\epsilon_i$), so that σ_i is composed of two factors: ϵ_i, the error inherent in p_i alone, and δ_i, the magnification of that error due to the presence of the other parameters. Unlike σ and ϵ, δ is not influenced by noise level. It reflects only the interdependence of the parameters within the current experimental design. A dependency value of, for example, ≥100 indicates either that we have overparameterized the model (e.g., too many site classes) or that our data range is somehow insufficient to resolve the parameters. The dependency value δ is the square root of the variance inflation factor described in the statistical literature (3).

In addition to *p*, *se*, and *dep*, the fitted function is provided for graphic comparison and goodness-of-fit tests, some of which are provided in the program but are not detailed here.

References

1. Moler, C., S. Herskovitz, J. Little, and S. Bangert. *MATLAB User's Guide*. The MathWorks, Inc., Natick, MA (1987).
2. Rodbard, D., and G. H. Weiss. Mathematical theory of immunoradiometric (labeled antibody) assays. *Anal. Biochem.* **52**:10–44 (1973).
3. Ketelslegers, J. M., G. D. Knott, and K. J. Catt. Kinetics of gonadotropin binding by receptors of the rat testis: analysis by a nonlinear curve-fitting method. *Biochemistry* **14**:3075–3083 (1975).
4. Karlsson, M. O., and A. Neil. Estimation of ligand binding parameters by simultaneous fitting of association and dissociation data: a Monte Carlo simulation study. *Mol. Pharmacol.* **25**:1–9 (1989).
5. Gear, C. W. The automatic integration of ordinary differential equations. *Commun. Assoc. Comput. Machinery* **14**:176–179 (1971).
6. Bard, Y. *Nonlinear Parameter Estimation*. Academic Press, New York (1974).
7. Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* **2**:164–168 (1944).
8. Marquardt, D. W. An algorithm for least-squares determination of nonlinear parameters. *J. Assoc. Comput. Machinery* **II**:431–441 (1963).
9. Belsley, D. A., E. Kuh, and R. E. Welsch. *Regression Diagnostics*. Wiley, New York (1980).
10. Draper, N. R., and H. Smith. *Applied Regression Analysis*. Wiley, New York (1966).

11. Oliva, D., F. Pocchiari, L. Allievi, G. E. Rovati, and S. Nicosia. Non-serotonergic ^{3}H-ketanserin binding sites in human platelets: characteristics and interaction with calcium antagonists. *Pharmacol. Res.* **26**:187–199 (1992).
12. Metters, K. M., N. Sawyer, and D. W. Nicholson. Microsomal glutathione S-transferase is the predominant leukotriene C_4 binding site in cellular membranes. *J. Biol. Chem.* **269**:12816–12823 (1994).
13. Rovati, G. E., D. Rodbard, and P. J. Munson. DESIGN: computerized optimization of experimental design for estimating K_d and B_{max} in ligand binding experiments. I. Homologous and heterologous binding to one or two classes of sites. *Anal. Biochem.* **174**:636–649 (1988).
14. Rodbard, D., P. J. Munson, and A. De Lean. Improved curve-fitting, parallelism testing, characterization of sensitivity and specificity, validation, and optimization for radioligand assays, in *Proceedings of Symposium on Radioimmunoassay and Related Procedures in Medicine*, Vol. 1. International Atomic Energy Agency, Vienna, Austria, 469–504 (1978).
15. Bates, D. M., and D. G. Watts. *Nonlinear Regression Analysis and Its Applications*. Wiley, New York, 53 (1988).

Send reprint requests to: Dr. G. Enrico Rovati, Institute of Pharmacological Sciences, University of Milan, Via Balzaretti 9, 20133 Milan, Italy. E-mail: rovati@isfunix.farma.unimi.it