Inversion of Markov Processes to Determine Rate Constants from Single-Channel Data

Meyer B. Jackson
Department of Physiology, University of Wisconsin Medical School, Madison, Wisconsin 53706 USA

ABSTRACT The determination of rate constants from single-channel data can be very difficult, in part because the single-channel lifetime distributions commonly analyzed by experimenters often have a complicated mathematical relation to the channel gating mechanism. The standard treatment of channel gating as a Markov process leads to the prediction that lifetime distributions are exponential functions. As the number of states of a channel gating scheme increases, the number of exponential terms in the lifetime distribution increases, and the weights and decay constants of the lifetime distributions become progressively more complicated functions of the underlying rate constants. In the present study a mathematical strategy for inverting these functions is introduced in order to determine rate constants from single-channel lifetime distributions. This inversion is easy for channel gating schemes with two or fewer states of a given conductance, so the present study focuses on schemes with more states. The procedure is to derive explicit equations relating the parameters of the lifetime distribution to the rate constants of the scheme. Such equations can be derived using the equality between symmetric functions of eigenvalues of a matrix and sums over principle minors, as well as expressions for the moments, derivatives, and weights of a lifetime distribution. The rate constants are then obtained as roots to this system of equations. For a gating scheme with three sequential closed states and a single gateway state, exact analytical expressions were found for each rate constant in terms of the parameters of the three-exponential closed-time distribution. For several other gating schemes, systems of equations were found that could be solved numerically to obtain the rate constants. Lifetime distributions were shown to specify a unique set of real rate constants in sequential gating schemes with up to five closed or five open states. For kinetic schemes with multiple gating pathways, the analysis of simulated data revealed multiple solutions. These multiple solutions could be distinguished by examining two-dimensional probability density functions. The utility of the methods introduced here are demonstrated by analyzing published data on nicotinic acetylcholine receptors, $GABA_A$ receptors, and NMDA receptors.

INTRODUCTION

The kinetic analysis of gating transitions of single ion channels has yielded a wealth of information about the conformational dynamics of these proteins. Most kinetic analysis of channel gating is based on the assumption of a Markov process, in which a small number of closed and open states interconvert with specified rates. Using well established mathematical methods one can derive the lifetime distributions for virtually any model of interest (Colquhoun and Hawkes, 1981, 1995; Fredkin et al., 1985), and these lifetime distributions serve as a fundamental theoretical prediction in the analysis of experimental data. Single-channel lifetime distributions generally take the form of a sum of exponentials, and the best-fitting function can be found with a number of readily available computer programs (French and Wonderlin, 1992; Jackson, 1992; Colquhoun and Sigworth, 1995; Heinemann, 1995). Unfortunately, once the fitting is accomplished it is very difficult to use these results to determine the rate constants of the underlying kinetic model. The problem is that although it is straightforward to go from a model to the predicted lifetime distribution, it is more difficult to perform the reverse operation and determine the underlying model and rate constants from the parameters characterizing the lifetime distributions. This inversion problem is of considerable practical importance because the primary goal of most kinetic studies of ion channels is to identify the best model and determine the rate constants.

For simple kinetic models (i.e., models with no more than two states aggregated into the same conductance) this inverse problem is straightforward. However, as models are made more complicated by adding more states, the determination of the rate constants from experimental lifetime distributions becomes more difficult. In fact, the majority of ion channels investigated are governed by multi-state kinetic schemes of sufficient complexity to make the analysis and interpretation of single-channel data a fairly involved enterprise. With channels governed by such complicated schemes, investigators are rarely able to make full use of lifetime distributions, partly because of the difficulty of the mathematical inversion problem. In these cases rate constants can be estimated without constructing lifetime distributions by using the method of likelihood maximization (Horn and Lange, 1983; Blatz and Magleby, 1986; Colquhoun and Sigworth, 1995; Qin et al., 1996). These methods are very powerful, but the widespread use of lifetime distributions in the analysis of single-channel kinetics

Received for publication 18 February 1997 and in final form 19 May 1997.

Address reprint requests to Meyer Jackson, Department of Physiology, SMI 129, University of Wisconsin Medical School, 1300 University Ave., Madison, WI 53706. Tel.: 608-262-9111; Fax: 608-265-5512. E-mail: mjackson@macc.wisc.edu.

0006-3495/97/09/1382/13 $2.00

provides a strong incentive to improve methods of extracting kinetic information from this form of data.

Another problem in the analysis of single-channel kinetics is that the data often cannot be used to identify a unique kinetic scheme. The basic connectivity of states as well as the rate constants are then indeterminate (Fredkin et al., 1985; Bauer et al., 1987; Kienker, 1989). The criterion for identifiability is based on whether the number of rate constants in the model is equal to the number of independent quantities in the lifetime distributions. For models satisfying this condition the theoretical studies cited above define a mathematics problem that can be solved to determine the rate constants, as well as evaluate the uniqueness of solutions. The present study is of this mathematics problem. This analysis has led to a new method for determining rate constants from single-channel data. With the equations derived here, a very small amount of code in the popular computer program MATHCAD (Windows version 5.0; Mathsoft, Cambridge, MA) renders quantitative estimates of rate constants directly from lifetime distributions for moderately complicated kinetic models of channel gating. The performance of these equations was evaluated, and data from published studies were used to illustrate the application of these methods.

THEORY

General method

Channel gating is treated as a discrete-state continuous-time aggregated Markov process. Here, the focus is on the lifetime distribution of a particular aggregate of channel states (e.g., the aggregate of all closed states or of all open states). We define a column vector, p, whose elements are the probabilities of being in the different states of the aggregate, subject to the condition of remaining in the aggregate. This vector obeys a differential equation of the following form (Colquhoun and Hawkes, 1995).

$$\frac{dp}{dt} = \mathbf{Q}_a p \tag{1}$$

The square matrix $\mathbf{Q}_a$ is composed of interconversion probabilities between every pair of states within the aggregate, together with the probabilities of exiting transitions out of the aggregate. $\mathbf{Q}_a$ can also be seen as a partition taken from the larger $\mathbf{Q}$ matrix that would appear in the complete Chapman-Kolmogoroff equation for the full kinetic model (Cox and Miller, 1965). The aggregate is denoted in Eq. 1 by the subscript a. When treating a closed-state aggregate, the subscript will be c, and it will be understood that p is restricted to probabilities of being in different closed states.

Equation 1 can be solved to give $p(t)$ (Colquhoun and Hawkes, 1981, 1995; Fredkin et al., 1985), and the sum of the elements of $p(t)$ gives the lifetime distribution of the aggregate (the open-time or closed-time distribution) as a sum of exponentials. The decay constants are the characteristic values or eigenvalues of $\mathbf{Q}_a$. It can be shown that the weights of the exponentials, X_i, are given by the expression

$$X_i = (\mathbf{W}^{-1}p^o)_i\left(\sum_j \mathbf{W}_{ij}\right) \tag{2}$$

where $\mathbf{W}$ is a matrix in which each column is an eigenvector of $\mathbf{Q}_a$. p^o denotes the probability vector at the start of a residence in the aggregated state for which the lifetime probability distribution is to be calculated; p^o is determined from the stationary occupancy probabilities of the model (Colquhoun and Hawkes, 1981, 1995).

If a channel gating scheme has n closed states and m open states, the open-time and closed-time distributions will be sums of n and m exponentials, respectively. Adding up the total number of decay constants and weights gives $2(n + m - 1)$ quantities to use in determining the rate constants of a model (Fredkin et al., 1985). (The weights of a lifetime distribution are constrained to add up to one, so the number of usable weights is one fewer than the number of exponentials.) Fredkin et al. (1985) pointed out that if there are no loops in the gating scheme, then the number of rate constants will also be equal to $2(n + m - 1)$, raising the possibility that the decay constants and weights of the lifetime distributions can be used to solve for the rate constants. The question addressed here is how to do that. The approach taken is to derive equations explicitly relating the rate constants of the kinetic scheme to the decay constants and weights of the lifetime distributions. These constraints are then solved to find the rate constants. In this way the Markov process is inverted.

The constraints used here are of four basic kinds. First, a general set of relations is exploited between the jth order symmetric function of the eigenvalues of a matrix and the sum over the determinants of all $j \times j$ principal minors (Horn and Johnson, 1985). These relations derive from the fact that the characteristic polynomial of a matrix is invariant with respect to a similarity transformation. Fredkin et al. (1985) proposed using some of these relations in the analysis of single-channel data. The highest order symmetric function of the eigenvalues is the product of all of them, and this is equal to the determinant of the matrix. The first order symmetric function is the sum of the eigenvalues, and this is equal to the trace of the matrix. An $n \times n$ square matrix will have n such relationships, each of which can be used as a constraint. These constraints are mathematically equivalent to using explicit formulas for the eigenvalues as roots of the characteristic equation. However, the eigenvalues cannot be written explicitly for matrices 5×5 and larger, and for 3×3 and 4×4 matrices the expressions are complicated.

Another class of constraints can be derived as moments of the lifetime distribution. The kth moment has the follow-

ing general form.

$$\langle t_a^k \rangle = k(-1)^k \sum_{j=1}^{n} (-\mathbf{Q}_a^{-k} p^o)_j = k(-1)^k \sum_{j=1}^{n} \frac{X_j}{\lambda_j^k} \qquad (3)$$

The sum in the middle is over the elements of the vector obtained as the product of a matrix and a vector, and the sum on the right is over all of the exponential terms of the lifetime distribution. Since the dimension of $\mathbf{Q}_a$ equals the number of exponentials, these sums are over the same number of terms, n. Equation 3 can be derived by multiplying Eq. 1 by $t^k \mathbf{Q}_a^{-1}$ and integrating by parts k times. Most of the uses of this equation in the present study will be with $k = 1$ to give the mean lifetime.

A third class of constraints is obtained from derivatives of the lifetime distribution at $t = 0$. Since the sum of the elements of p gives the lifetime distribution, a sum over elements in Eq. 1 gives the slope of the lifetime distribution at $t = 0$ as $\Sigma \mathbf{Q}_a p^o$. Equation 1 can be differentiated k times with respect to time. The terms of the resulting vector can then be summed on each side to produce a set of constraints of the form

$$\sum_{j=1}^{n} (\mathbf{Q}_a^k p^o)_j = \sum_{j=1}^{n} \lambda_j^k X_j \qquad (4)$$

This equation with $k = 1$ provides a very useful expression when the aggregate has only a single state into which gating transitions arrive. In this case the initial condition specifies one element of p^o as one, corresponding to the gateway state, and all the other elements of p^o as zero. Regardless of the manner in which the various states within an aggregate interconvert, Eq. 4 with $k = 1$ will give the following expression for a model with a single gateway state.

$$\sum \text{over all exit rates} = -\sum_{j=1}^{n} \lambda_j X_j \qquad (5)$$

This follows because when p^o contains one element equal to one and zeros everywhere else, the product $\mathbf{Q}_a p^o$ will be the single column of $\mathbf{Q}_a$ corresponding to the element of p^o that is one. The terms of the column of $\mathbf{Q}_a$ corresponding to transitions within the aggregate sum to zero to conserve probability (and satisfy detailed balance). However, the exit transitions do not conserve probability, and so the sum of the elements of that column will retain only the exit rates. Since all of the exit rates are associated with the gateway state, the sum of elements of $\mathbf{Q}_a p^o$ will produce the left side of Eq. 5. This point can be seen by inspection of the matrices in Eqs. 8 and 19 given below. The rightmost column of each sums to α. Equation 5 is a useful formula because of its general applicability to models with a single gateway state.

For kinetic schemes with multiple gateway states, some of the same rate constants appear in both the open-time and closed-time distributions. Explicit expressions can then be derived for the weights of a lifetime distribution. This gives a fourth class of constraints, which will become clearer in the examples below.

There is no obvious strategy for selecting which of the above constraints to use for a particular model other than simplicity and ease of solution. The inversion of Markov processes will now be illustrated with several examples.

Model I: CCCO

Consider a linear scheme with three closed states, one of which can open to the only open state.

$$C_1 \underset{d_1}{\overset{k_1}{\rightleftarrows}} C_2 \underset{d_2}{\overset{k_2}{\rightleftarrows}} C_3 \underset{\beta}{\overset{\alpha}{\rightleftarrows}} O \qquad (6)$$

Because there are three closed states, the closed-time distribution predicted by this model is a sum of three exponentials

$$P_c = X_1 e^{\lambda_1 t} + X_2 e^{\lambda_2 t} + X_3 e^{\lambda_3 t} \qquad (7)$$

The decay constants, λ_1, λ_2, and λ_3 are the eigenvalues of $\mathbf{Q}_c$, which for this model has the following form.

$$\mathbf{Q}_c = \begin{pmatrix} -k_1 & d_1 & 0 \\ k_1 & -d_1 - k_2 & d_2 \\ 0 & k_2 & -d_2 - \alpha \end{pmatrix} \qquad (8)$$

All closures begin in state C_3 to give the initial distribution of closed states as the vector $p^o = (0,0,1)$. The closed-time distribution is determined by five rate constants, k_1, k_2, d_1, d_2, and α. The channel closing rate, β, is the decay constant of the single-exponential open-time distribution, and this rate constant does not appear in the closed-time distribution.

To solve for the five rate constants, we need five constraints. Equation 5 gives the slope of the closed-time distribution at $t = 0$ as

$$\alpha = -\lambda_1 X_1 - \lambda_2 X_2 - \lambda_3 X_3 \qquad (9)$$

Thus, one of the rate constants, α, is determined and this simplifies the remaining problem to finding four constraints to solve for four unknowns.

The equivalence between sums over determinants of principal minors and symmetric functions of eigenvalues provides three constraints. From the determinant of $\mathbf{Q}_c$ we have

$$k_1 k_2 \alpha = -\lambda_1 \lambda_2 \lambda_3 = a_1 \qquad (10)$$

The symbol a_1 was introduced for $\lambda_1 \lambda_2 \lambda_3$ to facilitate analysis below. Similar symbols will be introduced in analogous situations. For the trace of $\mathbf{Q}_c$ we have

$$k_1 + k_2 + d_1 + d_2 + \alpha = -\lambda_1 - \lambda_2 - \lambda_3 = a_2 \qquad (11)$$

The sum over 2×2 principal minors of $\mathbf{Q}_c$ gives a third constraint.

$$k_1 k_2 + k_1 d_2 + k_1 \alpha + k_2 \alpha + d_1 d_2 + d_1 \alpha = \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \lambda_2 \lambda_3 = a_3 \qquad (12)$$

These last three equations can also be found in Colquhoun and Hawkes (1981). The final constraint needed to complete this problem was derived by considering the mean lifetime of the closed state. Setting $k = 1$ in Eq. 3 yields.

$$\langle t_c \rangle = \frac{k_1k_2 + k_1d_2 + d_1d_2}{\alpha k_1k_2} = -\frac{X_1}{\lambda_1} - \frac{X_2}{\lambda_2} - \frac{X_3}{\lambda_3} = a_4 \quad (13)$$

Equations 10–13 form a system of four equations that can be used to solve for the four unknowns, k_1, k_2, d_1, and d_2. Using the MAPLE symbolic processor of MATHCAD yielded explicit expressions for the rate constants, k_1, k_2, d_1, and d_2, in terms of α (from Eq. 9), a_1, a_2, a_3, and a_4. For k_1 and d_2 we have

$$k_1 = -\frac{a_1(a_1a_4 - \alpha^2 + \alpha a_2 - a_3)}{a_1^2a_4^2 + \alpha a_1a_2a_4 - 2a_1a_3a_4 + \alpha^2a_3 - \alpha a_2a_3 + a_3^2 - \alpha a_1} \quad (14)$$

and

$$d_2 = \frac{a_1a_4 - \alpha^2 + \alpha a_2 - a_3}{\alpha} \quad (15)$$

Rather than give the expressions for k_2 and d_1 (the expression for d_1 is lengthy), note that with k_1 determined from Eq. 14, and α from Eq. 9, one can then use Eq. 10 to obtain k_2. The only remaining unknown parameter is then d_1, which can be determined from either Eq. 11, 12, or 13.

An alternative solution to this same system was obtained by using the constraint based on the second derivative of the closed-time probability distribution at $t = 0$ (Eq. 4 with $k = 2$). This gave

$$\alpha^2 + d_2\alpha = \lambda_1^2X_1 + \lambda_2^2X_2 + \lambda_3^2X_3 \quad (16)$$

With α determined from Eq. 9, d_2 can be solved for. Any three of the four Eqs. 10–13 can then be selected to solve for the remaining rate constants, k_1, k_2, and d_1. Yet another solution to the inversion of the CCCO model can be found in Yang (1989).

The expressions for the rate constants obtained by either method contain no square or higher order roots. This is significant because it means that there are no multiple or imaginary solutions. Thus, a point in the five-dimensional space defined by the parameters of the closed-time distribution maps to a unique and real point in the five-dimensional space of rate constants in the CCCO model. It is more difficult to determine whether all of the rate constants must be positive. If a negative rate constant were obtained, then this would not be physically real; a negative rate constant makes no sense. During extensive searches with arbitrarily selected values, no combination of λ and X values could be found that gave negative rate constants. The rate constants were always positive, suggesting that every possible three exponential closed-time distribution with negative decay constants, and with weights adding up to one, corresponds to a unique and physically possible version of the CCCO model.

Model II: CCCOO

The CCCO model can be made more complicated by adding more states. An additional open state can be added in three different places, connecting to any one of the three closed states. The first example to be examined here is the CCCOO model, where both of the open states are reached from C_3.

$$C_1 \underset{d_1}{\overset{k_1}{\rightleftarrows}} C_2 \underset{d_2}{\overset{k_2}{\rightleftarrows}} C_3 \underset{\beta}{\overset{\alpha}{\rightleftarrows}} O_1 \qquad C_3 \underset{\gamma}{\overset{\delta}{\rightleftarrows}} O_2 \quad (17)$$

The analysis of this model is identical with that of the CCCO model above, except that α is replaced everywhere by $\alpha + \gamma$. The values of α and γ can then be determined from the relative frequencies of opening into O_1 and O_2. It is worth noting that if transitions are permitted directly between these two open states, then α and γ cannot be uniquely specified from the experimentally observed lifetime distributions, because of the indeterminacy of models containing loops (Fredkin et al., 1985; Bauer et al., 1987; Kienker, 1989). There will be three constraints from the double exponential open-time distribution, along with a constraint from detailed balance. The sum $\alpha + \gamma$ can still be determined from the right side of Eq. 9 (see Eq. 5), but with five constraints and six unknowns there will be infinitely many solutions for the rate constants involving open states, leaving these quantities indeterminate. Regardless of these issues, the formulas derived above for the CCCO model (Eqs. 9 through 16) can still be used to determine the interconversion rates between closed states.

Model III: CCCCO

Extending the linear sequence of closed states to four and retaining a single open state gives the CCCCO model.

$$C_1 \underset{d_1}{\overset{k_1}{\rightleftarrows}} C_2 \underset{d_2}{\overset{k_2}{\rightleftarrows}} C_3 \underset{d_3}{\overset{k_3}{\rightleftarrows}} C_4 \underset{\beta}{\overset{\alpha}{\rightleftarrows}} O \quad (18)$$

Seven of the eight rate constants of this model influence the closed-time distribution, namely k_1, k_2, k_3, d_1, d_2, d_3, and α. The closed-time distribution will be a sum of four exponentials, providing seven independent quantities (four decay constants and three weights) for formulating seven constraints. For this model

$$\mathbf{Q}_c = \begin{pmatrix} -k_1 & d_1 & 0 & 0 \\ k_1 & -d_1 - k_2 & d_2 & 0 \\ 0 & k_2 & -d_2 - k_3 & d_3 \\ 0 & 0 & k_3 & -d_3 - \alpha \end{pmatrix} \quad (19)$$

and $p^o = (0, 0, 0, 1)$. As with the CCCO model, Eq. 5 gives the opening rate, α.

$$\alpha = -\lambda_1X_1 - \lambda_2X_2 - \lambda_3X_3 - \lambda_4X_4 \quad (20)$$

This reduces the problem to setting up six constraints to solve for six unknowns. The equalities between symmetric

functions of eigenvalues and sums of determinants of principal minors give the following four constraints. The trace gives

$$k_1 + k_2 + k_3 + d_1 + d_2 + d_3 + \alpha = -\lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 \quad (21)$$

The 2×2 principal minors give

$$\begin{aligned} &d_1\alpha + d_1d_3 + k_2d_3 + k_2\alpha + d_1k_3 + d_1d_2 + k_3\alpha + d_2d_3 \\ &\quad + d_2\alpha + k_1k_2 + k_2k_3 + k_1k_3 + k_1d_2 + k_1d_3 + k_1\alpha \quad (22) \\ &\quad = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\lambda_4 + \lambda_2\lambda_3 + \lambda_2\lambda_4 + \lambda_3\lambda_4 \end{aligned}$$

The 3×3 principal minors give

$$\begin{aligned} &d_1k_3\alpha + d_1d_2d_3 + d_1d_2\alpha + k_2k_3\alpha + k_1k_2d_3 \\ &\quad + k_1k_2\alpha + k_1k_3\alpha + k_1d_2d_3 + k_1d_2\alpha + k_1k_2k_3 \quad (23) \\ &\quad = \lambda_1\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_4 + \lambda_1\lambda_3\lambda_4 + \lambda_2\lambda_3\lambda_4 \end{aligned}$$

The determinant gives

$$\alpha k_1k_2k_3 = \lambda_1\lambda_2\lambda_3\lambda_4 \quad (24)$$

A fifth constraint was obtained from the mean closed time using Eq. 3 with $k = 1$.

$$\begin{aligned} \langle t_c \rangle &= \frac{d_1d_2d_3 + k_1d_2d_3 + k_1k_2d_3 + k_1k_2k_3}{\alpha k_1k_2k_3} \\ &= -\frac{X_1}{\lambda_1} - \frac{X_2}{\lambda_2} - \frac{X_3}{\lambda_3} - \frac{X_4}{\lambda_4} \end{aligned} \quad (25)$$

In a first effort to find a sixth constraint, the second moment of the closed-time distribution was tried (from Eq. 3 with $k = 2$). The analytical expression obtained with the symbolic processor of MATHCAD was too long to reproduce easily here. Furthermore, MATHCAD did not find roots for this system very well. Fortunately, another approach worked better. Eq. 4 with $k = 2$ yielded the following result.

$$\alpha^2 + d_3\alpha = \lambda_1^2X_1 + \lambda_2^2X_2 + \lambda_3^2X_3 + \lambda_4^2X_4 \quad (26)$$

which is very similar to Eq. 16, except that d_3 replaces d_2. This equation was used to solve for d_3. The five equations above (Eqs. 21–25) were then used to solve for the five remaining rate constants, k_1, k_2, k_3, d_1, and d_2. The numerical root solver of MATHCAD performed well with this system.

Although analytic expressions were not obtained for the rate constants, analysis with the aid of Eq. 4 showed that any closed-time distribution of the CCCCO model specifies a unique set of real rate constants. This can be seen as follows. Equation 20 specifies a single real value for α, and Eq. 26 specifies a single real value for d_3. With $k = 3$, Eq. 4 gives the following

$$\begin{aligned} &-k_3d_3\alpha - d_3^2\alpha - 2d_3\alpha^2 - \alpha^3 \\ &\quad = \lambda_1^3X_1 + \lambda_2^3X_2 + \lambda_3^3X_3 + \lambda_4^3X_4 \end{aligned} \quad (27)$$

With values for d_3 and α already in hand, this equation specifies a unique and real value for k_3. Successive derivatives are increasingly complicated, but the pattern continues. The fourth derivative determines a unique real value for d_2 and the fifth derivative determines a unique real value for k_2. Equation 24 can then be used to determine k_1 uniquely from α, k_2, and k_3, and with all of the other parameters determined, Eq. 21 can be used to calculate a unique value for d_1. Thus, the CCCCO and CCCO models share a fundamental mathematical property in that any closed-time distribution with the appropriate number of exponentials specifies a unique set of real rate constants.

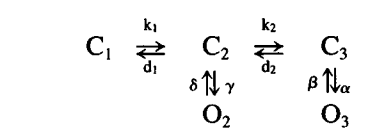
This approach of using derivatives from Eq. 4 was useful in terms of establishing an important mathematical property of the model, but successive derivatives should become harder and harder to distinguish from the quantity $\lambda_1^kX_1$ (where λ_1 is the fastest decay constant), and this is likely to make these equations less useful from a practical standpoint of solving for the rate constants.

Longer chains

Application of Eq. 4 to $\mathbf{Q}_c$ for a model with five sequential closed states showed that a new parameter could be determined with each successive derivative in the sequence α, d_4, k_4, d_3, k_3, etc. Based on the pattern established for these three models, it is reasonable to conjecture that linear models of arbitrary length specify a real one-to-one map between the parameters of the lifetime distribution and the rate constants of the gating scheme. The same analysis can be applied to linear sequences of open states as well. Thus, as long as a channel follows a linear sequence of state transitions, this conjecture for uniqueness of solutions can be extended to any purely sequential gating mechanism.

Model IV: CCOCO

Consider a model with an additional open state that can be reached only by the gating of the closed state C_2.



(28)

As with the CCCO model, the closed-time distribution is still a sum of three exponentials, but in the CCOCO model the closed-time distribution depends on six rate constants rather than five. The additional rate constant is γ, the rate of C_2 opening. This means that the five experimentally determined parameters of the closed-time distribution are not enough to specify the rate constants. However, this model does not contain a loop, and the rate constants can be determined. Because of the multiple gateway states in this model the open-time and closed-time distributions depend on some of the same rate constants. Thus, a sixth constraint can be derived from the open-time distribution of the

CCOCO model, which will be a sum of two exponentials

$$P_o = Y_1 e^{\kappa_1 t} + Y_2 e^{\kappa_2 t} \tag{29}$$

The relative frequency of opening into O_2 and O_3 can be calculated from the stationary probabilities of closed state occupancy to give the following constraint.

$$\frac{\alpha k_2}{\gamma d_2} = \frac{Y_1}{Y_2} \tag{30}$$

The other constraints from the CCCO model can be used if suitably modified. $\mathbf{Q}_c$ from Eq. 8 is altered only by the addition of γ to $\mathbf{Q}_{c2,2}$, and p^o is $[0, \gamma d_2/(\gamma d_2 + \alpha k_2), \alpha k_2/(\gamma d_2 + \alpha k_2)]$. The mean closed time (from Eq. 3 with $k = 1$) gives the following constraint

$$\frac{d_1 d_2 + k_1 d_2 + k_1 k_2}{k_1(\alpha k_2 + \gamma d_2)} = -\frac{X_1}{\lambda_1} - \frac{X_2}{\lambda_2} - \frac{X_3}{\lambda_3} \tag{31}$$

The slope of the closed-time distribution at $t = 0$ (Eq. 4, $k = 1$) gives

$$\frac{\alpha^2 k_2 + \gamma^2 d_2}{\alpha k_2 + \gamma d_2} = -\lambda_1 X_1 - \lambda_2 X_2 - \lambda_3 X_3 \tag{32}$$

Given $\mathbf{Q}_c$ for the present model, Eqs. 10–12 of the CCCO model become the following

$$k_1(k_2\alpha + \gamma d_2 + \gamma\alpha) = -\lambda_1\lambda_2\lambda_3 \tag{33}$$

$$k_1 + k_2 + d_1 + d_2 + \gamma + \alpha = -\lambda_1 - \lambda_2 - \lambda_3 \tag{34}$$

$$(k_1 + \alpha)(k_2 + \gamma) + (k_1 + d_1)(d_2 + \alpha) + \gamma d_2 = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \tag{35}$$

Equations 30–35 represent a set of six constraints that can be used to solve for the six unknowns, k_1, k_2, d_1, d_2, γ, and α of the CCOCO model in Scheme 18.

Unlike the constraints set up for the CCCO model, Eqs. 30–35 do not yield an analytical solution. Nor could analytical solutions be found for any of the other models analyzed below. However, the above system of constraints yielded numerical solutions with the aid of MATHCAD. The analysis of real and simulated data with these equations will be discussed below, after additional models are analyzed.

Model IV: COCCO

In this model the most distant closed states can open.

$$\begin{array}{ccccc} C_1 & \underset{d_1}{\overset{k_1}{\rightleftarrows}} & C_2 & \underset{d_2}{\overset{k_2}{\rightleftarrows}} & C_3 \\ {\scriptstyle\delta}\uparrow\downarrow{\scriptstyle\gamma} & & & & {\scriptstyle\beta}\uparrow\downarrow{\scriptstyle\alpha} \\ O_1 & & & & O_3 \end{array} \tag{36}$$

$\mathbf{Q}_c$ and p^o for this model were set up and analyzed by methods identical to those used above for the CCOCO model. The following six constraints were derived.

$$\frac{\alpha k_1 k_2}{\alpha k_1 k_2} = \frac{Y_1}{Y_2} \tag{37}$$

$$\frac{d_1 d_2 + k_1 d_2 + k_1 k_2}{\alpha k_1 k_2 + \gamma d_1 d_2} = -\frac{X_1}{\lambda_1} - \frac{X_2}{\lambda_2} - \frac{X_3}{\lambda_3} \tag{38}$$

$$\frac{\alpha^2 k_1 k_2 + \gamma^2 d_1 d_2}{\alpha k_1 k_2 + \gamma d_1 d_2} = -\lambda_1 X_1 - \lambda_2 X_2 - \lambda_3 X_3 \tag{39}$$

$$k_1 k_2 \alpha + \gamma d_1 d_2 + \gamma\alpha d_1 + \gamma\alpha k_2 = -\lambda_1\lambda_2\lambda_3 \tag{40}$$

$$k_1 + k_2 + d_1 + d_2 + \gamma + \alpha = -\lambda_1 - \lambda_2 - \lambda_3 \tag{41}$$

$$k_1 k_2 + \gamma d_1 + \gamma k_2 + d_1 d_2 + d_1\alpha + k_2\alpha + k_1 d_2 + k_1\alpha + \gamma d_2 + \alpha\gamma = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \tag{42}$$

Each is respectively analogous to Eqs. 30 through 35 above.

Model V: COCOCO

Allowing all three closed states to open gives a model for which both open times and closed times are distributed as a sum of three exponentials.

$$\begin{array}{ccccc} C_1 & \underset{d_1}{\overset{k_1}{\rightleftarrows}} & C_2 & \underset{d_2}{\overset{k_2}{\rightleftarrows}} & C_3 \\ {\scriptstyle\phi}\uparrow\downarrow{\scriptstyle\varepsilon} & & {\scriptstyle\delta}\uparrow\downarrow{\scriptstyle\gamma} & & {\scriptstyle\beta}\uparrow\downarrow{\scriptstyle\alpha} \\ O_1 & & O_2 & & O_3 \end{array} \tag{43}$$

Taken together, the experimentally determined open-time and closed-time distributions yield a total of 10 constraints, which can be used to determine the 10 rate constants of the model. The three closing rates are trivially the three decay constants of the open-time distribution (it would not be as simple if interconversions between the open states were allowed). The closed-time distribution will depend on seven of the rate constants of the COCOCO model, the two k values, the two d values, and the three opening rates, α, γ, and ϵ. Thus, we need seven constraints.

For the determinant of $\mathbf{Q}_c$ we have

$$k_1 k_2 \alpha + k_1 d_2 \gamma + k_1\alpha\gamma + d_1 d_2 \varepsilon + d_1\alpha\varepsilon + k_2\alpha\varepsilon + d_2\gamma\varepsilon + \alpha\gamma\varepsilon = -\lambda_1\lambda_2\lambda_3 \tag{44}$$

For the trace we have

$$k_1 + k_2 + d_1 + d_2 + \alpha + \gamma + \varepsilon = -\lambda_1 - \lambda_2 - \lambda_3 \tag{45}$$

For the 2×2 principal minors we have

$$(k_1 + \varepsilon)(k_2 + \gamma) + (k_1 + \varepsilon)(d_2 + \alpha) + (d_1 + \gamma)(d_2 + \alpha) + k_2\alpha + d_1\varepsilon = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \tag{46}$$

The initial slope from Eq. 4 with $k = 1$ gives

$$\frac{\alpha^2 k_1 k_2 + \gamma^2 k_1 d_2 + \varepsilon^2 d_1 d_2}{\alpha k_1 k_2 + \gamma k_1 d_2 + \varepsilon d_1 d_2} = -\lambda_1 X_1 - \lambda_2 X_2 - \lambda_3 X_3 \quad (47)$$

For the mean closed time we have (by Eq. 3)

$$\frac{k_1 k_2 + k_1 d_2 + d_1 d_2}{\alpha k_1 k_2 + \gamma k_1 d_2 + \varepsilon d_1 d_2} = -\frac{X_1}{\lambda_1} - \frac{X_2}{\lambda_2} - \frac{X_3}{\lambda_3} \quad (48)$$

Finally, the weights of the open-time distribution give these two constraints

$$\frac{\alpha k_1 k_2}{\alpha k_1 k_2 + \gamma k_1 d_2 + \varepsilon d_1 d_2} = \frac{Y_1}{Y_1 + Y_2 + Y_3} \quad (49)$$

$$\frac{\gamma k_1 d_2}{\alpha k_1 k_2 + \gamma k_1 d_2 + \varepsilon d_1 d_2} = \frac{Y_2}{Y_1 + Y_2 + Y_3} \quad (50)$$

Equations 44–50 thus constitute seven constraints that can be used to solve for the seven rate constants in the COCOCO model.

ERRORS IN PARAMETERS

The time constants and weights from lifetime distributions used in the foregoing analysis have confidence limits determined by the errors of instrumental noise and statistical analysis. It is important to know how the errors in these quantities carry over to the rate constants determined by Markov process inversion. Most exponential fitting software gives a covariance matrix $V(x_i, x_j)$ to aid in evaluating errors. In this case the variables x_i and x_j are any pair of the $2n-1$ parameters, either weights or decay constants, from an n-exponential lifetime distribution. If the rate constants are linear functions of these parameters within an interval comparable to the confidence limits of the parameters, then the variances of the rate constants can be determined by the following double sum over all pairs of indices (Hald, 1952).

$$V(\phi) = \sum_{ij}^{2n-1} \frac{\partial\phi \partial\phi}{\partial x_i \partial x_j} V(x_i, x_j) \quad (51)$$

where the symbol ϕ denotes any rate constant determined from Markov process inversion. Thus, the key quantities in determining how errors are propagated during Markov process inversion are the derivatives of the rate constants with respect to the lifetime distribution parameters. These derivatives can be determined numerically by varying the input parameters and repeating the inversion process. Some of the equations given in the analysis of models above can be differentiated to provide analytical expressions for the derivatives (e.g., Eq. 5 and Eq. 11). Equation 51 should be useful in evaluating the errors in parameters from a single experiment. However, for Markov process inversion as with any other method of analysis, the final assessment of error should be based on an evaluation of the variability between replicate experiments. Equation 51 will be used to assess the propagation of errors during the analysis of data from the acetylcholine receptor below.

IMPLEMENTATION

For all of the above models the computer program MATHCAD (Windows version 5.0, Mathsoft, Cambridge, MA) was used to determine the rate constants numerically as roots to the corresponding set of constraints. All computing was performed on a personal computer with a pentium or cyrix 6x86 CPU. Although the analytic solution to the CCCO model could be used, it was more convenient to handle all of the models with the same protocol based on numerical solution.

Each root solving procedure employed the following format. 1) The data were entered. This included the weights and decay constants of the closed-time distribution. For multiple gateway models parameters from the open-time distribution were also entered. 2) The λ and X values were used to calculate the constraining parameters such as mean closed time, initial slope, symmetric functions of eigenvalues, etc. For the CCCO and CCCCO models the opening rate, α, was calculated directly according to Eqs. 9 and 20, respectively, and for the CCCCO model d_3 was calculated with Eq. 26. 3) Initial guesses for the rate constants were entered. 4) The MATHCAD "Minerr" function was used to solve for the rate constants using the appropriate constraints. These constraints were Eqs. 10–13 for the CCCO model (Scheme 6), Eqs. 21–25 for the CCCCO model (Scheme 18), Eqs. 30–35 for the CCOCO model (Scheme 28), Eqs. 37–42 for the COCCO model (Scheme 36), and Eqs. 44–50 for the COCOCO model (Scheme 43). 5) The rate constants thus obtained were checked by computing the lifetime distributions from the solution to Eq. 1. 6) Finally, an error statistic was determined for each constraint. Each deviation from equality was evaluated, normalized to the constraining value determined from the experimental quantities, and squared.

These MATHCAD root solving protocols found the rate constants in times ranging from less than a second for the CCCO model to up to about three seconds for the COCOCO model. For the more complicated models the root solvers were often unable to find a solution for a given set of initial guesses, but trial and error with different initial guesses eventually lead to a solution. It was important to check the λ and X values as well as the error. The Minerr function in MATHCAD often found a local minimum rather than a point with zero error.

TESTS

To test the root solving procedures described above, simulated open-time and closed-time distributions were generated by solving Eq. 1 for different choices of rate constants. $\mathbf{Q}_c$ was set up for each model tested and the eigenvalues and eigenvectors were determined numerically with MATH-

CAD. The weights were calculated according to Eq. 2. These simulated data were entered into the root solver for the relevant model to see how well the original rate constants were recovered. Because single-channel analysis becomes more difficult with closely spaced time constants and small weights, the rate constants were selected to give time constants differing by more than a factor of two, and weights $>\sim 0.05$. Otherwise, rate constants were varied arbitrarily to generate simulated lifetime distributions with parameters spanning as wide a range as possible.

In nearly all of the tests conducted, the root solver returned rate constants equal to those used to generate the simulated data set. In a few cases where the recovered rates deviated from the original values, this could be attributed to the fact that when some of the λ values were small their product was small enough to compromise computational accuracy. This was more of a problem with the CCCCO model, where the product of four small numbers (Eq. 24) could be quite small. The problem was remedied by scaling all the λ values up by a factor of 10, recovering the rate constants, and dividing back down by 10, to yield rate constants equal to those used to generate the simulated lifetime distribution.

Because the CCCO model was solved analytically and tested against published data described in Applications below, there was little need to test it with simulations. However, with kinetic schemes containing multiple gateway states the situation was complicated, and tests with simulated data yielded some valuable insights.

Multiple gateway schemes

Ten sets of simulated lifetime distributions were generated and used to test the CCOCO model. In all cases the root solver returned the original rate constants, although it was often necessary to try five or more sets of initial guesses of rate constants before a solution was obtained. For one set of simulated data the solution was found only when the initial guesses were very close to the correct rate constants.

In six of the 10 data sets simulated for the CCOCO model, alternative solutions were found. The root solver found rate constants different from the original set used to generate the closed-time and open-time distributions. This indicates that in contrast to sequential gating schemes, the inversion operation for the CCOCO model gives multiple solutions. Although different models with the same number of open and closed states can be fitted to the same one-dimensional lifetime distributions (Magleby and Weiss, 1990), the present results indicate that even for one model there can be multiple solutions.

In four of the simulated data sets one additional solution was found, in one set two additional solutions were found, and in one set three additional solutions were found, for a total of four. Four of the 10 simulated data sets yielded only one solution, despite extensive variation of initial guesses. However, the possibility cannot be ruled out that one or more additional solutions eluded detection. Table 1 shows the four sets of rate constants that all produce the same lifetime distributions. These results show that in spite of the equality between the number of rate constants and the number of quantities in the lifetime distributions (Fredkin et al., 1985), the rate constants are not uniquely specified by one-dimensional lifetime distributions.

It should be noted that sometimes when an input data set was deliberately varied from a simulated set, no solution was found, in spite of extensive variation of the initial guesses. This may mean that there are some one-dimensional lifetime distributions that are inconsistent with the CCOCO model. Thus, the absence of real and positive roots could provide a useful criterion for eliminating models.

Tests of the COCCO and COCOCO models gave similar results. Three data sets were simulated from the COCCO model and two data sets were simulated from the COCOCO model. Original rate constants were recovered and multiple solutions were found.

The existence of multiple sets of rate constants that generate the same closed-time and open-time distributions raises the question of whether these alternative sets of rate constants are fundamentally indistinguishable with single-

TABLE 1 Multiple solutions

Lifetime Distribution					
$\lambda_1 = -1.532$	$\lambda_2 = -0.6066$	$\lambda_3 = -0.2707$	$X_1 = 0.1198$	$X_2 = 0.1616$	$Y_2/Y_1 = 0.333$
	Rate Constants				
	k_1	k_2	d_1	d_2	α	γ
Set 1 (original)	0.5	0.05	0.05	0.5	1.0	0.3
Set 2	0.5	0.3968	0.2143	0.5	0.2333	0.5556
Set 3	1.064	0.1914	0.2193	0.1751	0.1753	0.5749
Set 4	0.8205	0.0403	0.4631	0.1228	0.4803	0.4733

This table lists the solutions found by the root solver for the CCOCO model using the simulated lifetime distribution parameters on the top line. λ and X refer to the decay constants and weights from the closed-time distribution, and Y_2/Y_1 refers to the ratio of weights in the open-time distribution (Eq. 29). The original set of rate constants used to simulate the lifetime distribution was recovered as a solution (set 1). In addition, three other solutions were recovered for different choices of initial guesses (sets 2–4). All four of these sets of rate constants generate a lifetime distribution with the weights and decay constants on the top line. However, the weights in the two-dimensional probability density function (Eq. 52) generated by each set were different.

TABLE 2 Rate constants for the nicotinic receptor

[ACh] μM	k_1 s^{-1}	k_2 s^{-1}	d_1 s^{-1}	d_2 s^{-1}	α s^{-1}
10	315	1510	142	38900	43700
30	1250	3850	162	37200	44300
100	6360	13400	183	38700	45500
300	18300	43300	1100	32400	46600
From Table 3 of Sine et al. (1990).	6×10^7 M^{-1} s^{-1}	1×10^8 M^{-1} s^{-1}	250	40000	45000

Entering the λ and X values read from Fig. 9 of Sine et al. (1990), the CCCO solver returned the above values for rate constants. The rate constants are those indicated in Scheme 6. The bottom line of this table contains values for rate constants that Sine et al. determined from their data and reported in Table 3 of their paper. Note that k_1 and k_2 determined here are simple velocities in units of s^{-1} and those of Sine et al. are association rates in units of M^{-1} s^{-1}. See Fig. 1 for a plot of k_1 and k_2 versus acetylcholine concentration. The association rates determined here are given in the legend of Fig. 1, and are in good agreement with those of Sine et al. (1990).

channel data. Since it has been shown that two-dimensional lifetime probability density functions delimit the information contained in single-channel data (Fredkin et al., 1985; Bauer et al., 1987), this question can be addressed by calculating and comparing these functions. $f(t_c, t_o)dt_c dt_o$, the probability density function for observing a closed time of duration t_c followed by an open time of duration t_o, was derived for the CCOCO model as

$$f(t_c, t_o)dt_c dt_o = (\gamma\delta p_2(t_c)e^{-\delta t_o} + \alpha\beta p_3(t_c)e^{-\beta t_c})dt_c dt_o \quad (52)$$

where $p_2(t_c)$ and $p_3(t_c)$ are the triple exponential solutions to Eq. 1 for the CCOCO model. Since the solving procedure makes use of a complete set of constraints based on principal minors and symmetric functions of eigenvalues, the eigenvalues of $\mathbf{Q}_c$ are the same, and the decay constants of the six exponentials in Eq. 52 are the same. However, the weights of these exponentials in Eq. 52 are different for the four different sets of rate constants in Table 1. Thus, the four solutions found by analyzing the one-dimensional probabilities in this example can be distinguished by analyzing lifetime correlations in single-channel data. The inspection of two-dimensional probability density functions should be a useful approach in situations where multiple solutions arise, and this will be relevant to the analysis of data from $GABA_A$ receptor channels below.

APPLICATIONS

The nicotinic receptor

The CCCO model was used by Sine et al., (1990) to interpret their data on the *Torpedo* nicotinic acetylcholine receptor channel. The three closed states are assumed to be unligated, monoligated, and diligated receptors (Scheme 6). Sine et al. (1990) determined the rate constants by fitting this model to closed-time distributions by likelihood maximization, using the explicit roots of the cubic characteristic equation to calculate the likelihood function. They also fitted closed-time distributions to a sum of three exponentials and plotted the decay constants and weights as a function of acetylcholine concentration. Reading the numbers from their plot (Fig. 9 of their paper, corresponding to results at 22°C in normal Ca^{2+}), and using the root solver protocol described above for the CCCO model, I obtained the estimates for the rate constants shown in Table 2. The values reported in Table 3 of Sine et al., (1990) are reproduced here in Table 2 for comparison. d_1, d_2, and α are in fairly good agreement. Table 3 of Sine et al. showed a grand scheme without dividing data up on the basis of concentration. It is especially significant that the quantities d_1, d_2, and α of Table 2 are very similar for 10, 30, and 100 μM acetylcholine. This is expected because these rate processes do not reflect association reactions. Only at 300 μM acetylcholine are these numbers different, and this probably reflects open channel block by the neurotransmitter at the highest concentration. This would compromise the validity of the CCCO model by adding an additional state to the scheme.

The values obtained for k_1 and k_2 increase linearly with acetylcholine concentration (Fig. 1), as expected for an association process. Lines fitted these plots very well, and passed quite close to the origin. The slopes gave association rates of $0.62 \pm 0.02 \times 10^8$ M^{-1} s^{-1} and $1.45 \pm 0.03 \times 10^8$ M^{-1} s^{-1}, which again are similar to the values of Sine et al. (1990) (reproduced here at the bottom of Table 2). The differences between d_1 and d_2 and between k_1 and k_2 were noted by Sine et al. (1990) as showing nonequivalence of the two agonist binding sites of the receptor, and this nonequivalence is also evident from the concentration de-

TABLE 3 Propagation of errors

	$d \ln \alpha$	$d \ln k_1$	$d \ln k_2$	$d \ln d_1$	$d \ln d_2$
$d \ln \lambda_1$	0.98	−0.004	0.026	−0.046	1.07
$d \ln \lambda_2$	0.019	0.14	0.84	2.8	−0.063
$d \ln \lambda_3$	0.0029	0.86	0.13	−1.8	−0.0096
$d \ln X_1$	−0.53	−0.44	−2.7	−1.1	0.97
$d \ln X_2$	0.011	−0.43	0.48	−2.1	−0.035

Changes in the values of the input parameters (λ and X) resulted in changes in the rate constants computed for the CCCO model. These derivatives provide an assessment of error propagation during Markov process inversion (Eq. 51). The derivatives were determined numerically. Each entry gives a measure of the sensitivity of the estimated rate constant to error in the input. The derivatives of the logarithms of the input parameters with respect to the logarithms of the output parameters were used to make each change relative. This table was generated with the 30 μM acetylcholine data of Sine et al. (1990).

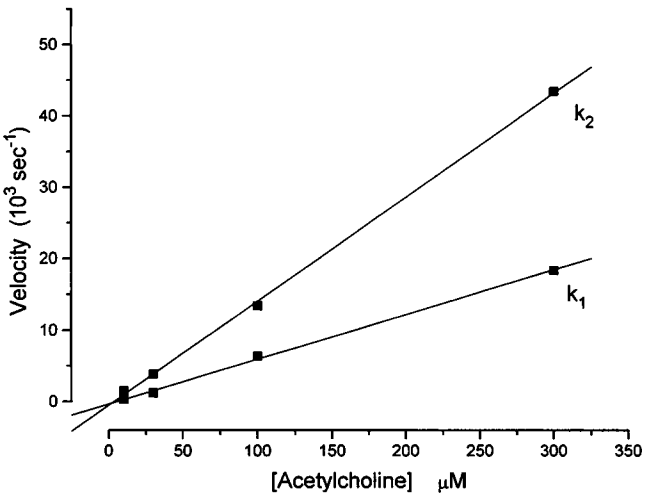


FIGURE 1 Plots of k_1 and k_2 versus acetylcholine concentration. Rate constants were determined by applying the CCCO root solver to the closed-time distribution data of Sine et al. (1990). The steps corresponding to these rate constants are indicated in Scheme 6. The values of the rate constants of Scheme 6 determined here are given in Table 2. This plot shows that k_1 and k_2 increase linearly with concentration, as expected for binding steps. The best fitting lines gave association rates of 0.62×10^8 M^{-1} s^{-1} and 1.45×10^8 M^{-1} s^{-1} as slopes, and these values are close to the values of Sine et al. (1990) (Table 2).

pendence of the frequency of channel opening of singly ligated and doubly ligated receptors (Jackson, 1988).

Errors in parameters

To evaluate the errors introduced by this method, the sum of squares error (as described in Implementation) was minimized while constraining one rate constant at a time to a fixed value. Using the 30 μM data of Sine et al. (1990) for this purpose, the error was plotted versus parameter value in Fig. 2. This plot shows that the error is very large except for a narrow range around the roots. Only for d_1 was an additional local minimum seen near $d_1 = 1$. However, there was no root at this point, as the expanded inset shows. Because the error increases very steeply for deviations from the correct solution, Fig. 2 suggests that rate constant estimation by this method will not increase the error very much beyond that incurred during the fitting of exponential functions to the lifetime distribution.

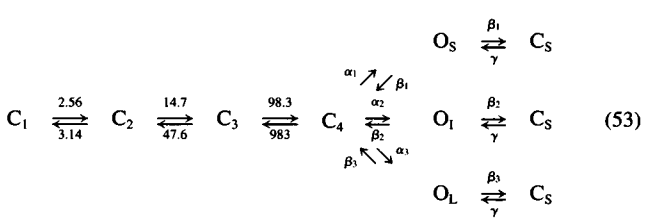
To evaluate how the error is propagated from the λ and X values to α and to the k and d values, the partial derivatives of the rate constants with respect to the experimental input parameters were evaluated numerically (Eq. 51) and shown in Table 3. These derivatives were normalized to the magnitudes of the parameters so that they can be read as the fractional change in one relative to the fractional change in the other. This shows that for the data set selected [30 μM acetylcholine from Sine et al. (1990)], only in 4 of the 25 connections between input and output was there a significant amplification of error, and this was by less than a factor of three. Thus, for the data set chosen here there was relatively little error introduced by the present method of determining rate constants.

The $GABA_A$ receptor

Weiss and Magleby (1989) found that the gating of the $GABA_A$ receptor channel can be described by a model like the CCOCO model (Scheme 28), but with a third open state connected to O_3. The closed-time distribution is still a sum of three exponentials, with the same functional dependence on the rate constants as the CCOCO model. To apply the CCOCO model root solver to the data of Weiss and Magleby it is only necessary to combine the weights of the two slowest components of the open-time distribution when calculating Y_2 in Eq. 30. The open-time and closed-time distributions were calculated from the rate constants for patch 1 of Table 2 in Weiss and Magleby (1989). When the root solver for the CCOCO model was given these data, it recovered the original set of rate constants and generated two additional solutions. The two-dimensional probability density function (Eq. 52) was determined for each of the three solutions and found to be different. Thus, as was seen with simulated data, the multiple solutions from a set of real data can be distinguished on the basis of correlations. The method of analysis used by Weiss and Magleby (1989) incorporated information about lifetime correlations, so their rate constants could not be one of the false solutions.

The NMDA receptor

Closed-time distributions of the NMDA receptor contain several exponential components, but it is thought that all openings arise from fully ligated receptors (Benveniste and Mayer, 1991; Clements and Westbrook, 1991). Kleckner and Pallotta (1995) used a "gateway model" to interpret single-channel data on the NMDA receptor. Their gateway model is the part of the following kinetic scheme indicated by the gray bar.



(53)

The model employed by Kleckner and Pallotta (1995) did not include the closed states on the left, but since fitting the closed-time distribution required five exponentials, we can combine the gateway model of Kleckner and Pallotta with a linear sequence of closed states shown in the model above (Scheme 53). The rates shown are in units of s^{-1} and were determined in the present study, as described below. Kleckner and Pallotta (1995) determined the values of the other rate constants indicated by Greek letters.

Since the state C_s is isolated from the other closed states we can subtract the component corresponding to this state

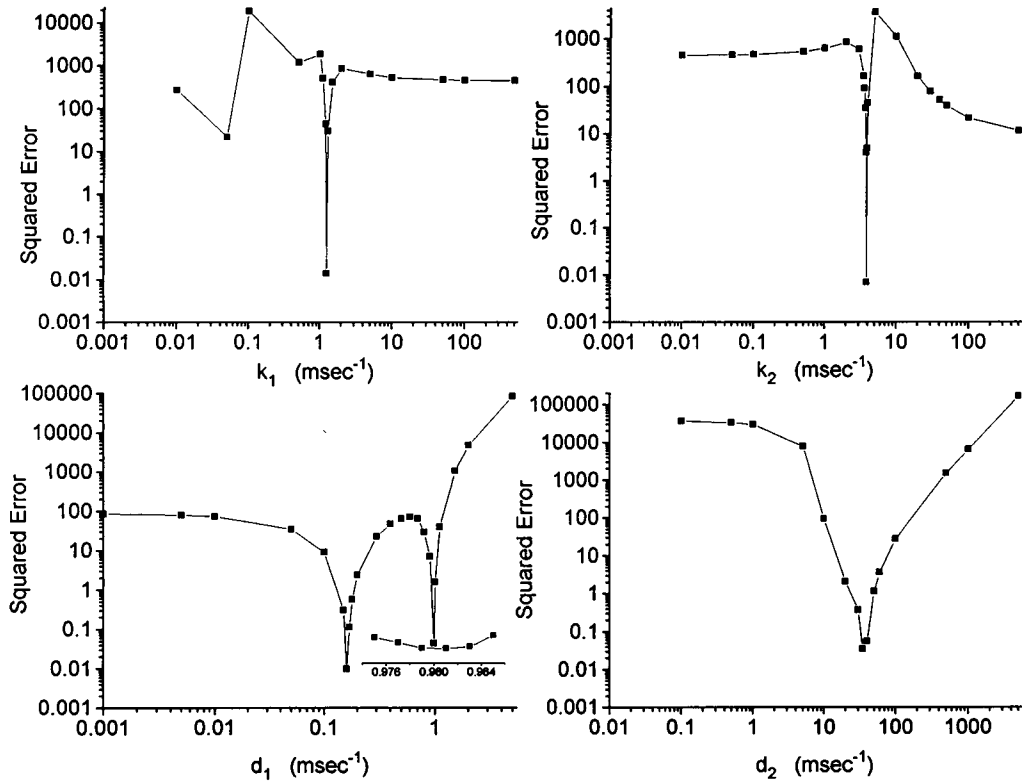


FIGURE 2 The squared error is plotted versus the value of a constrained parameter for the 30 μM data of Sine et al. (1990). Squared error was determined as the sum of squares of deviations from equality for each constraint, normalized to the constraining parameter (a_1, a_2, a_3, and a_4, in Eqs. 10–13). The error goes to zero only in a very narrow region around the root. Note that the lowest points plotted are not the roots. Zero error is obtained but this cannot be shown on a logarithmic scale. These plots show how this method of analysis yields well-defined parameters for this set of data. With d_1 an additional local minimum was found. This region was expanded to show that there was no root in the vicinity.

(the fastest component) from the closed-time distribution and use the CCCCO model to interpret the remaining four exponential components of closed times. I used the four slowest time constants and corresponding weights stated on page 414 of Kleckner and Pallotta (1995), and renormalizing the weights of these four components to add up to one. When the root solver for the CCCCO model was given this input it returned the rate constants shown in the above scheme. The sum of the three α values determined from Eq. 5 was 288 s^{-1}.

The rate constants determined here for interconversions between closed states may well correspond to association and dissociation of the agonist NMDA. Kleckner and Pallotta (1995) used 1 μM NMDA, which is well below saturating. The coligand, glycine, was present in saturating concentrations so that glycine binding is less likely to account for any of these steps. Two molecules of NMDA are thought to bind to gate the channel (Benveniste and Mayer, 1991; Clements and Westbrook, 1991), so it is difficult to decide which of the three closed-state interconversions shown are ligand association and dissociation events. However, regardless of the question of which steps are binding steps, the rate constants differ. The equilibrium dissociation constants calculated from these rates for 1 μM NMDA were 1.23, 3.2, and 10 μM for binding to states C_1, C_2, and C_3, respectively. Thus, as was found for the nicotinic acetylcholine receptor (Jackson, 1988; Sine et al., 1990), the binding sites of the NMDA receptor may not be equivalent.

DISCUSSION

This study introduces a new method for determining the rate constants for ion channel transitions. Tests on both simulated and experimental data indicated that Markov process inversion can be applied in the analysis of many commonly encountered forms of single-channel data. Because this analysis can be performed with about one page of code written with MATHCAD, a popular and widely used computer program, these methods should be accessible to researchers in the investigation of channel gating mechanisms. Several different models were analyzed here. While the mathematical strategy can in principle be adapted to other models, as these models increase in complexity the complexity of the mathematical analysis will also increase. It is not clear at what point the mathematical complexity, the number of rate constants and mathematical constraints, and the number of multiple solutions will become a serious impediment. For now, the method of Markov process inversion will most likely be useful for moderately complex models such as those analyzed here.

A major advantage claimed here for this new method of analysis is the ease of implementation with a standard computer program. However, the problem of multiple solutions places demands on the computer software that were not easily met by MATHCAD. Trying out many initial guesses and getting at most one solution at a time is not a very reliable way to find all of the possible solutions. More robust root solving software would be very useful to search parameter space thoroughly and return a complete set of all possible solutions. More powerful software may also be helpful in applying Markov process inversion to more elaborate models by making it easier to find the roots of more complicated systems of equations.

This work showed that some kinetic schemes have unique solutions and others have multiple solutions. A better understanding of the mathematics underlying these differences would be useful. Two **Q** matrices with the same eigenvalues must be related by a similarity transformation. Kienker (1989) used similarity transformations to analyze the distinguishability of models with different connectivities. It is likely that theoretical analysis along similar lines will provide a more rigorous basis for determining the existence and number of multiple solutions. This should also help in establishing methods of distinguishing between multiple solutions using higher dimensional (correlation) analysis. Multiple solutions in single-channel analysis have previously been found in relation to the problem of missed intervals (Blatz and Magleby, 1986). However, the multiple solutions found here are for simulated data, where the missed-interval problem is irrelevant.

Missed intervals are an important problem in the analysis of single-channel kinetics (Roux and Sauvé, 1985; Blatz and Magleby, 1986; Crouzy and Sigworth, 1990; Jackson, 1992), and this issue was neglected in the present study. For models with a single gating pathway the procedures described by Blatz and Magleby (1986) can be used to correct the lifetime distributions for missed events. Markov process inversion can be applied to these corrected lifetime distributions to obtain estimates of rate constants with an appropriate missed-interval correction. For more complicated gating schemes the missed interval problem is more of a challenge and correcting lifetime distributions will depend on the details of the gating scheme and the specific values of the rate constants. In these situations the new likelihood maximization methods of Qin et al. (1996) are clearly advantageous. It may also be possible to incorporate a strategy for missed event correction into the present inversion method. The **Q** matrix can be modifed to take missed events into account (Roux and Sauvé, 1985; Crouzy and Sigworth, 1990). The mathematical approaches of the present study could then be applied to this modified **Q** matrix to derive equations and solve for rate constants.

Much of the previous theoretical work in single-channel kinetics has been devoted to distinguishing between different models for channel gating on the basis of state number and connectivity (Horn and Vandenberg, 1984; Fredkin et al., 1985; Colquhoun and Hawkes, 1987; Blatz and Magleby, 1989; Jackson, 1992). The tests here on simulated data showed that when lifetime distributions were deliberately altered from those simulated by the CCOCO model, the root solver for that model was sometimes unable to find a new solution. If this is in fact because no real solution exists, then this would mean that finding or not finding solutions could serve as a criterion for evaluating models with different connectivities. However, since there are examples of successful fits of one-dimensional distributions to models with different connectivities (Magleby and Weiss, 1990), this criterion can only be useful in negating models. The appropriate behavior of rate constants for candidate models may also help in evaluating the connectivity of a channel gating mechanism. For example, the increase in the rate of a binding step with drug concentration illustrates this approach (Fig. 1) and confirms the CCCO model for this channel.

Making rigorous distinctions between different models with the same numbers of open and closed states requires some form of analysis of correlations between event lifetimes. As presented here, the method of Markov process inversion deals only with the one-dimensional distribution functions and ignores the correlation information vital to distinguishing between models. However, Markov process inversion has the potential to simplify the analysis of lifetime correlations. The rate constants determined by Markov process inversion for different candidate models can be used to calculate correlation coefficients (e.g., by integration of Eq. 52). These can then be compared with the experimental correlation coefficients. This approach may also be useful in choosing between the multiple solutions obtained for models with multiple gating pathways. However, in applying these approaches to real data a full analysis of the errors of parameters and of predicted correlation coefficients will be essential.

Kinetic processes described by a system of linear differential equations in the form of Eq. 1 are very common in biophysics. Such equations describe not only stationary systems such as those analyzed here, but also nonstationary systems studied with voltage jump and other rapid perturbation methods. Markov process inversion may be applicable to these other forms of kinetic data as well. The relationships between symmetric functions of eigenvalues and sums over principal minors does not depend on the particular form of data, and should be readily adapted to other forms of experimental data including nonstationary data. In nonstationary experiments one often sees a maximum or minimum in the time course of a signal as channels inactivate or desensitize. At that point a time derivative is zero, and when that condition is imposed on Eq. 1, new constraints between the parameters can be derived. The derivatives at $t = 0$ (Eq. 4) also should be adaptable to other forms of kinetic data by using the appropriate initial state vector, p°. The mean lifetimes, moments of lifetime distributions, and explicit expressions for the weights of a distribution are specialized to stationary single-channel data, but if other strategies for deriving constraints can be tailored

to a specific type of experiment, the method of Markov process inversion will be useful in other kinetic applications as well as single-channel analysis.

Note added in proof: MATHCAD files for implementing the analysis described here will be made available at the following web site http://www.neuroscience.wisc.edu/faculty/jackson.html.

I thank Shyue-Fang Hsu for helpful discussions.

This research was supported by National Institutes of Health Grant NS23512.

REFERENCES

Bauer, R. J., B. F. Bowman, and J. L. Kenyon. 1987. Theory of the kinetic analysis of patch-clamp data. *Biophys. J.* 52:961–978.

Benveniste, M. J., and M. M. Mayer. 1991. Kinetic analysis of antagonist action at *N*-methyl-D-aspartic acid receptors: two binding sites each for glutamate and glycine. *Biophys. J.* 59:560–573.

Blatz, A. L., and K. L. Magleby. 1986. Correcting single channel data for missed events. *Biophys. J.* 49:967–980.

Blatz, A. L., and K. L. Magleby. 1989. Adjacent interval analysis distinguishes among gating mechanisms for the fast chloride channel from rat skeletal muscle. *J. Physiol.* 410:561–585.

Clements, J. D., and G. L. Westbrook. 1991. Activation kinetics reveal the number of glutamate and glycine binding sites on the NMDA receptor. *Neuron.* 7:605–613.

Colquhoun, D., and A. G. Hawkes. 1981. On the stochastic properties of single ion channels. *Proc. R. Soc. Lond.* 211:205–235.

Colquhoun, D., and A. G. Hawkes. 1987. A note on correlations in single ion channel records. *Proc. R. Soc. Lond. B.* 230:15–52.

Colquhoun, D., and A. G. Hawkes. 1995. A Q-matrix cookbook. *In* Single-Channel Recording. B. Sakmann and E. Neher, editors. Plenum Publishing Corp., New York. 589–633.

Colquhoun, D., and F. Sigworth. 1995. Fitting and statistical analysis of single channel records. *In* Single-Channel Recording. B. Sakmann and E. Neher, editors. Plenum Publishing Corp., New York. 483–587.

Cox, D. R., and H. D. Miller. 1965. The Theory of Stochastic Processes. Chapman and Hall, London.

Crouzy, S. C., and F. J. Sigworth. 1990. Yet another approach to the dwell time omission problem of single-channel analysis. *Biophys. J.* 58: 731–743.

Fredkin, D. R., M. Montal, and J. A. Rice. 1985. Identification of aggregated Markovian models: application to the nicotinic acetylcholine receptor. *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer.* Vol. 1. L. M. Le Carn and R. A. Olshen, editors. Wadsworth, Monterey, CA. 269–289.

French, R. J., and W. F. Wonderlin. 1992. Software for acquisition and analysis of ion channel data: choices, tasks, and strategies. *Methods Enzymol.* 207:711–728.

Hald, A. 1952. Statistical Theory with Engineering Applications. John Wiley & Sons, New York.

Heinemann, S. H. 1995. Guide to data acquisition and analysis. *In* Single-Channel Recording, B. Sakmann and E. Neher, editors. Plenum Publishing Corp., New York. 53–91.

Horn, R. A., and C. A. Johnson. 1985. Matrix Analysis. Cambridge University Press, New York.

Horn, R., and K. Lange. 1983. Estimating kinetic constants from single channel data. *Biophys. J.* 43:207–223.

Horn, R., and C. A. Vandenberg. 1984. Statistical properties of single sodium channels. *J. Gen. Physiol.* 84:505–534.

Jackson, M. B. 1988. Dependence of acetylcholine receptor channel kinetics on agonist concentration in cultured mouse muscle fibres. *J. Physiol.* 397:555–583.

Jackson, M. B. 1992. Stationary single channel analysis. *Methods Enzymol.* 207:729–746.

Kienker, P. 1989. Equivalence of aggregated Markov models of ion-channel gating. *Proc. R. Soc. Lond. B.* 236:269–309.

Kleckner, N. W., and B. S. Pallotta. 1995. Burst kinetics of single NMDA receptor currents in cell-attached patches from rat brain cortical neurons in culture. *J. Physiol.* 486:411–426.

Magleby, K. L., and D. S. Weiss. 1990. Identifying kinetic gating mechanisms for ion channels by using two-dimensional distributions of simulated dwell times. *Proc. R. Soc. Lond. B.* 241:220–228.

Qin, F., A. Auerbach, and F. Sachs. 1996. Estimating single-channel kinetic parameters from idealized patch-clamp data containing missed events. *Biophys. J.* 70:264–280.

Roux, B., and R. Sauvé. 1985. A general solution to the time interval omission problem applied to single channel analysis. *Biophys. J.* 48: 149–158.

Sine, S. M., T. Claudio, and F. Sigworth. 1990. Activation of *Torpedo* acetylcholine receptors expressed in mouse fibroblasts. *J. Gen. Physiol.* 96:395–437.

Weiss, D. S., and K. L. Magleby. 1989. Gating scheme for single GABA-activated Cl^- channels determined from stability plots, dwell-time distributions, and adjacent-interval durations. *J. Neurosci.* 9:1314–1324.

Yang, X. C. 1989. Characterization of stretch activated channels in Xenopus oocytes. Ph.D dissertation State University of New York at Buffalo.