Corrections

APPLIED MATHEMATICS, BIOPHYSICS. For the article “Using independent open-to-closed transitions to simplify aggregated Markov models of ion channel gating kinetics,” by William J. Bruno, Jin Yang, and John E. Pearson, which appeared in issue 18, May 3, 2005, of Proc. Natl. Acad. Sci. USA (102, 6326–6331; first published April 20, 2005; 10.1073/pnas.0409110102), in Fig. 3b, the minus signs have been omitted from the values of \( N_c - R \) and \( N_c - R \), due to a printer’s error. The corrected figure and its legend appear below. The conclusions are unchanged by this typographical error.

**Fig. 3.** Comparison of transition matrices for canonical forms. (a) BKU “uncoupled” form. (b) MIR form. White space denotes regions of the matrix filled with zeros. Shaded regions have nonzero values. Diagonal lines indicate a single diagonal strip of nonzero values. In b, the width of the gray regions is equal to the rank \( R \). When \( R \) takes its maximum value, the number of nonzero values is the same in the two forms.

BIOPHYSICS. For the article “Detection and characterization of the product of hydroethidine and intracellular superoxide by HPLC and limitations of fluorescence,” by Hongtao Zhao, Jöy Joseph, Henry M. Fales, Edward A. Sokoloski, Rodney L. Levine, Jeanette Vasquez-Vivar, and B. Kalyanaraman, which appeared in issue 16, April 19, 2005, of Proc. Natl. Acad. Sci. USA (102, 5727–5732; first published April 11, 2005; 10.1073/pnas.0501719102), the authors note that the last sentence of the Abstract, “We conclude that the HPLC/fluorescence assay using HE as a probe is more suitable reactive oxygen species for detecting intracellular \( O_2^- \),” should read “We conclude that the HPLC/fluorescence assay using HE as a probe is more suitable for detecting intracellular \( O_2^- \).” This change does not alter the conclusions of the article.

CELL BIOLOGY. For the article “A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization,” by Xiaolan Zhao and Günter Blobel, which appeared in issue 13, March 29, 2005, of Proc. Natl. Acad. Sci. USA (102, 4777–4782; first published February 28, 2005; 10.1073/pnas.0500537102), the authors note that on page 4778, in the second full paragraph in the right column, line 2, “Qri2” should read “Qri2/Nse4.” In line 12 of the same paragraph, “Nse4” should read “Nse5,” and in line 14, “Qir2” should read “Nse1, Nse3.” These changes do not affect the conclusions of the article.
Using independent open-to-closed transitions to simplify aggregated Markov models of ion channel gating kinetics

William J. Bruno*, Jin Yang*, and John E. Pearson†

Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, NM 87545

Edited by Harry L. Swinney, University of Texas, Austin, TX, and approved March 17, 2005 (received for review December 8, 2004)

Deducing plausible reaction schemes from single-channel current traces is time-consuming and difficult. The goal is to find the simplest scheme that fits the data, but there are many ways to connect even a small number of states (e.g., 2 million schemes with four open and four closed states). Many schemes make identical predictions. An exhaustive search over model space does not address the many equivalent schemes that will result. We have found a canonical form that can express all reaction schemes for binary channels. This form has the minimal number of rate constants for any rank (number of independent open–closed transitions), unlike other canonical forms such as the well-established “uncoupled” scheme. Because all of the interconduction transitions in the new form are independent, we refer to it as the manifest interconduction rank (MIR) form. In the case of four open and four closed states, there are four MIR form schemes, corresponding to ranks 1–4. For many models proposed in the literature for specific ion channels, the equivalent MIR form has dramatically fewer links than the uncoupled form. By using the MIR form we prove that all rank 1 topologies with a given number of open and closed states make identical predictions in steady state, thus narrowing the search space for simple models. Moreover, we prove that fitting to canonical form preserves detailed balance. We also propose an efficient hierarchical algorithm for searching for the simplest possible model consistent with a given data set.

On channels participate in many cell signaling pathways. They allow spatial and temporal control of calcium ions that regulate processes as diverse as fertilization, memory, muscle contraction (1), and immune response (2, 3). Abnormal channel function is implicated in pathologies such as cystic fibrosis (4), ataxia (5), and immune response (2, 3). Abnormal channel function is implicated in pathologies such as cystic fibrosis (4), ataxia (5), and immune response (2, 3). Abnormal channel function is implicated in pathologies such as cystic fibrosis (4), ataxia (5), and immune response (2, 3).

A Markov model of the dynamics of a channel consists of a reaction scheme or topology of allowed transitions between states, together with rate constants for these transitions. For the typical case of binary channels studied here, all open states have the same current, so the states can each be labeled by “O” for open or “C” for closed. The indistinguishability of the closed (or open) states from one another causes ambiguity in fitting models to data. Kienker (17) showed that the two topologies C–C–O and C–O–C are “equivalent”: for every set of rates for one there is a set of rates for the other such that both have identical steady-state behavior.

The space of topologies grows rapidly with the number of states. Our goal is to provide tools and strategies for navigating this space, allowing researchers to focus on a small, tractable set. The fitting of rate constants then can be carried out with existing ion channel analysis software, such as QUB (19), that is capable of handling any desired topology.

Methods

Here, we discuss the established concepts used in this work. Beyond these concepts, we allude to the mathematical techniques used after each result. Detailed proofs are contained in the supporting information, which is published on the PNAS web site. These proofs mainly depend on linear algebra, especially Kienker transformations (17), which are briefly reviewed there as well.

Most of the results presented here have exceptions which are, literally, infinitely uncommon. These exceptions happen when certain relationships among the rate constants cause degeneracies, such as two exactly identical decay rates in the open-time distribution. In the statements of the theorems below, those exceptions are indicated by the mathematical terms “almost all” or “almost every,” meaning “with infinitely rare exceptions.” Throughout, the term “topology” means the reaction scheme, usually displayed in a state transition diagram, without regard to reaction rates. “Model” means a Markov model with a particular topology and a particular set of reaction rates.

Equivalence and Canonical Forms. There are two related notions of equivalence: equivalent models and equivalent topologies. “Equivalent models” have identical steady-state dwell-time distributions. “Equivalent topologies” are capable of modeling the same range of steady-state behaviors, and therefore almost every model with one topology is equivalent to some model with the other (as in Definition 1 in the supporting information). Equivalent topologies will have the same number $N_O$ of open states and the same number $N_C$ of closed states, and in fact they must have the same number of “independent” O–C links (defined below). Beyond these constraints, it has so far been nontrivial to say

---

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: BIC, Bayesian Information Criterion; BKU, Bauer–Kienker uncoupled; DB, detailed balance; MIR, manifest interconduction rank.

*W.J.B., J.Y., and J.E.P. made significant contributions to this work.

†To whom correspondence should be addressed at: MS K-710, Los Alamos National Laboratory, Los Alamos, NM 87545. E-mail: pearson@lanl.gov.
whether two topologies are equivalent. Steady-state experiments can never distinguish among equivalent topologies.

Any collection of equivalent models may be viewed as ways in which different dynamics can generate the same steady-state statistics. Any prescription for expressing any set of equivalent entities in a unique way is called a “canonical form.” Representing a model in canonical form means finding an equivalent model with canonical topology, which should always be feasible for a good canonical form. The choice of canonical forms has previously been limited to just two, as discussed below, for which it was known how to find canonical models equivalent to almost any model.

Equivalent models are equivalent to a single model in canonical form, and this fact can be used to test whether two models are equivalent. Equivalent models will fit data equally well, regardless of the fitting technique (e.g., maximum likelihood or least squares). To obtain the best possible fit to steady-state data, it suffices to fit the data to models in canonical topologies, because models in other topologies are always equivalent to canonical models.

Known Canonical Forms. We will compare the canonical form introduced below to the canonical form described by Bauer et al. (12) and Kienker (17), which we refer to as Bauer–Kienker uncoupled (BKV) form. In BKV form, every open state is connected to every closed state and vice versa, but there are no open-to-open or closed-to-closed connections. This form has $2N_{o}N_{c}$ rate constants.

Physical models must have positive reaction rates. Negative reaction rates can lead to negative probabilities for individual states and consequently are difficult to interpret. Kienker noted (17) that models equivalent to a physical model can have negative rates. For example, it can be shown that any four-state model in BKV form must have negative rates to be equivalent to any physical model with topology $C$–$O$–$O$–$C$ (see $C$–$O$–$O$–$C$ Has Negative Rates in BKV Form in the supporting information). This result implies that negative rates should be allowed when fitting data to a canonical form. If the resulting canonical model is nonphysical, it can still be viewed as a parsimonious representation of the data using a generalized Markov process (20). Below, we show that negative rates can sometimes be avoided.

If the principle of detailed balance (DB) is not obeyed, BKV form models may require rates that are complex numbers to achieve the best fit. A generally applicable canonical form given by Larget (18) is similar to BKV form. Both have the same number of parameters, and both minimize the number of intracurrent reactions. In Larget’s form, complex numbers are never required, unlike in BKV form. Larget’s form always has negative rates, whereas the complex rates in BKV form can be removed by a slight change in the form that adds no more parameters (see Real and Complex Rates in the supporting information).

Dwell-Time Distributions. The dwell-time distribution $f_{t_{i,o}}(t_{c})$ is the probability density that the time interval for which the channel is open has duration $t_{o}$. Likewise, $f_{t_{i,c}}(t_{o}, t_{c})$ is the probability density of consecutive open and closed intervals with durations $t_{o}$ and $t_{c}$. Colquhoun et al. (8, 9) derived expressions for these distributions in terms of the “generator matrix,” $Q_{ij}$, which is the matrix of transition rates between the various states. Fredkin et al. (10, 11) showed that $f_{t_{o}}$ and $f_{t_{o}}$ are sufficient to fully characterize the steady-state data, under certain technical conditions that are almost always satisfied. All of the dwell-time distributions are sums of decaying exponentials. For example,

$$f_{t_{i,o}}(t_{c}) = \sum_{i=1}^{N_{o}} \sum_{j=1}^{N_{c}} \alpha_{ij} e^{-\lambda_{i}t_{o} - \omega_{ij}t_{c}},$$

where $\lambda_{i}$ and $\omega_{ij}$ are the open and closed decay constants, respectively, and $\alpha_{ij}$ is a matrix of amplitudes.

Identifiability. If different models make identical predictions about a particular class of experiments, the models are not “identifiable” with respect to those experiments. A topology is identifiable if it represents any changes to its rate constants result in changes to the steady-state dwell-time distributions. Conversely, topologies that have more parameters than there are degrees of freedom in the data are not identifiable. In this case, there is a continuum of rate constants for the topology that all yield identical statistics, and thus values for these rates cannot be uniquely determined from steady-state data, regardless of the quantity or quality of data. A trivial example of a nonidentifiable topology is one with two closed states and no open states. Obviously, patch-clamp current traces could not be used to infer the transition rates between the two states.

It has long been known (10) that there are at most $2N_{o}N_{c}$ degrees of freedom in steady-state data. Therefore, topologies with more than $N_{o}N_{c}$ links are not identifiable. The bound on the number of degrees of freedom can be improved (11) to $2R(N_{o} + N_{c} - R)$, where $R$ is the interconductance rank defined below.

Topologies with more than this many rates are not identifiable; topologies with fewer rates may or may not be identifiable. The problem of determining which topologies are identifiable remains unsolved (10).

Interconductance Rank. The canonical form that we introduce in this work makes explicit use of $R$, the interconductance rank. Before giving the physical interpretation of $R$, we review the mathematical definition. The rank of a model is defined as the matrix rank (number of independent rows or columns) of the matrix $Q_{oo}$, which is the block of $Q_{ij}$ representing open to closed transitions. We assume the ranks of $Q_{oo}$ and $Q_{co}$ are equal, which is generally true for reversible reactions. These ranks also equal the matrix rank of $a_{o}$, which can be estimated directly from the data without reference to a model (11). Therefore, equivalent models must have the same interconductance rank.

The rank of a topology is defined as the maximal rank of any model with that topology, and this rank equals the number of “independent” O–C links, where different independent links do not link to the same state; this equality is proved in Rank and Independent Reactions in the supporting information.

If a model is fit to data, the rank of the resulting model will equal either the rank of the data ($R_{o}$) or the rank of the topology of the model, whichever is smaller. For a model to have plausible generated data, the two ranks should be equal. If the topology used for the fitted model has a larger rank, then there will be linear dependencies among the rates that cannot be justified biochemically, unless there is a symmetry that can explain them.

The ranks of a few illustrative topologies are shown in Fig. 1. The rank of the topology in Fig. 1a is 1. The topologies shown in Fig. 1 b–d all have rank 2. Note that in Fig. 1d there are three Os connected to Cs and three Cs connected to Os, but the rank of the topology is only 2. Fredkin and Rice (11) noted that the lesser of the number of Os connected to Cs and the number of Cs
nonzero values is the same in the two forms. R filled with zeros. Shaded regions have nonzero values. Diagonal lines indicate necessary, and usually does not lead to plausible models. 

Fig. 3. Examples of reaction diagrams for canonical forms. (a) BKU canonical form for \( N_o = 6 \) and \( N_c = 5 \). The number of reversible reactions is \( N_oN_c - 30 \). (b) MIR canonical form for \( N_o = 6 \), \( N_c = 5 \), and \( R = 3 \). The number of reversible reactions is \( R(N_o + N_c - R) = 24 \).

connected to Os is an upper bound on a topology’s rank. Fig. 1d shows that this bound is not tight; note that no three of the O–C links are independent. Incidentally, this example is a subtopology (or special case) of BKU form, because it has no O–O or C–C links.

It can be shown (Theorem 8 in the supporting information) that when rank 1 data are fit to BKU form, the resulting number of nonzero rate parameters is \( 2N_oN_c \), far more than the \( 2(N_o + N_c - 1) \) necessary when \( R = 1 \). In this case the rank of the topology is the lesser of \( N_o \) and \( N_c \), whereas the rank of the model is 1. If \( N_o \) and \( N_c \) are both \( >1 \), then it is generally not plausible that rank 1 data were generated by a BKU form model; the model may fit the data, but there is no explanation for the linear dependencies of the rates. Thus, BKU form often obscures the rank, uses more links than necessary, and usually does not lead to plausible models.

Results

Manifest Interconductance Rank (MIR) Form. MIR form is defined by the following requirements: (i) all interconductance links are independent, and (ii) all intraconductance links involve at least one state with an interconductance link. The MIR form is illustrated in Figs. 2b and 3b. Fig. 2b shows how states can be arranged into four columns. The two central columns contain the same number of states, with Cs in one and Os in the other. In this case the states are linked in horizontal pairs to form the \( R \) interconductance O–C links. The remaining Cs and Os are in the first and last columns, respectively. Every C in the second column is connected to every other C, and Cs in the first column are not connected to each other. Connections for the Os follows the same pattern. The total number of reactions in MIR form is \( R(N_o + N_c - R) \), which gives the Fredkin bound of \( 2R(N_o + N_c - R) \) independent rate constants that can be recovered from steady-state data. The MIR form can be constructed for any rank up to the maximum possible rank (the lesser of \( N_o \) and \( N_c \)). Because all interconductance links are independent, the rank of the topology equals the number of O–C links. The maximum-rank MIR form has all lower rank MIR forms as subtopologies, making the full-rank MIR form as powerful as BKU form. A comparison of the two forms expressed as matrices of rate parameters is shown schematically in Fig. 3, in a case where the rank is not full.

There are two steps required to show that MIR form is canonical. The first step is to show that for almost any model, there is an equivalent model in canonical form and to give a prescription for how to find that model.

Theorem. Almost every model can be expressed in MIR form.

Mathematically, the MIR form corresponds to “diagonalizing” the interconductance transition matrices \( Q_{oc} \) and \( Q_{co} \). This transformation is almost always possible. The nondegeneracy conditions that comprise the exceptions are comparable to ones required for BKU form (see Manifest Interconductance Rank Form in the supporting information). The required diagonalization is performed by first diagonalizing the products \( Q_{oc}Q_{co} \) and \( Q_{co}Q_{oc} \). The mathematics of this procedure have been exploited previously in other fields such as canonical correlation analysis (21) and partial least-squares (22).

The theorems in Fig. 2 both have \( N_o = 5 \), \( N_c = 6 \). The rank of the MIR form topology in Fig. 2b is given by \( R = 3 \). The rank of the BKU form topology is fixed by \( R = \min(N_o, N_c) \), which is 5 in this case. In BKU form (Fig. 2a), there are 30 reversible transitions, whereas in MIR form (Fig. 2b), there are only 24. If one were to fit data generated by a rank-3 topology to BKU form, one could recover as many as 60 reaction rates (30 reactions), but these parameters would not all be independent. If one were to fit the same data to MIR form, one would recover at most 48 parameters (24 reactions), and these parameters would almost always be independent.

The second part of showing that MIR form is canonical is proving that its topology is identifiable. If this condition were not true, then expression of a model in MIR form would not be unique. We state our result as follows.

Theorem. MIR and BKU forms are identifiable for almost all parameters.

Proofs for the two forms are presented as Theorems 1 and 2 in the supporting information. Both are related to the fact that there is only one normalized similarity transformation, neglecting permutations, that can diagonalize a nondegenerate matrix.

Although generically MIR form has fewer links than BKU form, there are special topologies that have fewer links when expressed in BKU form. One example is the topology in Fig. 1d. This topology can be constructed by removing four links from the general BKU form, leaving the five links shown. Equivalent models in MIR form have eight links. Conversely, starting from MIR form and removing links results in topologies requiring more links when expressed in BKU form, regardless of rank. Thus, if models in nature tend to resemble simplified special cases of one form or the other, fitting to that form may result in simpler models, as the following examples will show.

Many channel models in the literature conform or nearly conform to MIR form. In Table 1, we count the number of MIR- and BKU-compatible reactions in four publications by different authors. Compatible reactions are those that are in the equivalent model in canonical form. A topology with all reactions compatible is already in the given canonical form. We shall see that when only a few reactions are not compatible, the canonical version of the model may have only a few more links than the original model.

Of the 20 models discussed in these articles, 7 are in MIR form, and none are in BKU form. Also, models in the literature tend to
Table 1. Resemblance of models proposed by different authors to BKU and MIR canonical forms

<table>
<thead>
<tr>
<th>Ref.</th>
<th>No. of models</th>
<th>BKU links, %</th>
<th>MIR links, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>11</td>
<td>31</td>
<td>76</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>36</td>
<td>80</td>
</tr>
<tr>
<td>30</td>
<td>2</td>
<td>47</td>
<td>93</td>
</tr>
<tr>
<td>31</td>
<td>1</td>
<td>29</td>
<td>93</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>33</td>
<td>79</td>
</tr>
</tbody>
</table>

have substantially more MIR-compatible than BKU-compatible links. Fig. 4 illustrates two of the cases where only one link was not compatible with MIR form.

Fig. 4 shows both proposed topologies, together with the MIR and BKU form topologies that can model the same data. These topologies are found by the transformation techniques described in Manifest Interconductance Rank Form and Kienker Transformations in the supporting information, and they will give the same results for almost any rate constants. In Fig. 4a, the resulting MIR form requires only two more links than the original topology to fit the same data. BKU form requires an additional five links as shown. Similarly, Fig. 4b shows a proposed topology with 14 links, only 1 of which is not MIR-form compatible. The MIR form topology has 15 links. The BKU form topology has 36 links.

In both examples in Fig. 4, when either canonical form is used, the number of links will be less than the $R(N_c + N_o - R)$ bound, because both topologies are sparsely linked to begin with. However, using MIR form results in significantly fewer links because these models are topologically close to being simplified special cases of MIR form.

The tendency for models in the literature to resemble MIR form more than BKU form may indicate a tacit belief among researchers that nonindependent O–C links are rare. This belief implies that negative rate constants may be required to find a canonical form model that is equivalent to some physical models. Rank 1 topologies do not suffer from this complication.

**Theorem.** Nonnegativity of rates is preserved when a rank 1 model satisfying DB is expressed in either MIR or BKU forms.

The proof (see Theorem 5 in the supporting information) relies on the special structure of these canonical forms in the rank 1 case. The proof is easy compared with the related result (28) that amplitudes $a_{ij}$ in the dwell-time distribution $f(t)$ (see Eq. 1) are positive if the open and closed dwell times are uncorrelated and DB is satisfied. Uncorrelated open and closed dwell times imply that $f_{ij}(t_o, t_c) = f(t_o)f(t_c)$, which is true for all rank 1 models (10). This fact is intuitive: if there is only a single gateway state linking the Os and Cs in a Markov model, there is no way to generate correlated open and closed times, because an individual Markov state has no memory. Note that the theorem does not apply to rank 2 topologies, such as the C–O–O–C example described earlier that does have negative rates in BKU form (but is already in MIR form).

The above result is not the only simplification that applies in the rank 1 case. We have proven for fixed $N_o$ and $N_c$ the following.

**Theorem.** All rank 1 topologies are equivalent.

This result is somewhat surprising. It does not hold for rank 2 (see None of C–O–O–C, O–C–O–C, and C–O–C–O are equivalent to each other in the supporting information). The rank 1 equivalence proof uses induction on the number of open and closed states (see Theorem 6 in the supporting information). The proof also can be adapted to show that given a rank 1 model with nonnegative rates obeying DB, every rank 1 topology has an equivalent model with nonnegative rates. This means that the “all rank 1 topologies are equivalent” result holds when restricting to physical models if DB

**Fig. 4.** Two topologies proposed in the literature and the topologies of the equivalent models in BKU and MIR forms. (a) A full rank model of the nicotinic acetylcholine receptor described in ref. 30. Of its eight links, seven are MIR-form compatible, whereas four are BKU-form compatible. Equivalent BKU form models have 13 links, compared with 10 links for MIR form. (b) A rank 4 model of the inositol trisphosphate receptor described in ref. 31. There are 14 links, 13 of which are compatible with MIR form, whereas only 4 are compatible with BKU form. As shown, equivalent MIR form models have 15 links, whereas equivalent BKU form models have 36 links.
Recent determination of atomic resolution structures of some ion channels have made possible molecular dynamics simulations of the opening or closing of these channels. These simulations are far too short (tens of nanoseconds) to compare with current traces, which describe events on the millisecond scale and longer. However, it could be possible to use numerical techniques to identify quasi-stable conformations of an ion channel that are candidates for the multiple open and closed states evidenced by patch-clamp experiments. To the extent that some of the relevant coordinates are preserved when the channel opens and closes, MIR form should allow simplification of the model, beyond any guaranteed simplification owing to the rank, as was seen in Table 1 and Fig. 4. If a ligand can be bound with the channel open or closed, then the ligand-binding coordinate would constitute an example of such preservation.

Atomic structures also can inform the choice of \( N_o \) and \( N_c \), based on the number of predicted ligand-binding sites, for example. Symmetries observed in the structure should also be reflected in the model, although the implications for the canonical forms are not clear. Molecular dynamics simulations also might shed light on the number of independent pathways between open and closed states, which should then be reflected in the rank \( R \) of the model.

**Exhaustive Search**

The brute force approach to exhaustive search is to search over all topologies. The optimization criterion should take into account the quality of the fit and the simplicity of the model; an established one that has been applied in this setting is the Bayesian Information Criterion (BIC) (15). The brute force approach has the advantage that only positive rates are needed to find the greatest likelihood. It has the disadvantage that there are \( 10^{10} \) trillion topologies for the case of 6 open and 6 closed states.

A more sophisticated exhaustive strategy would be to search over only one model in every class of equivalent models. This strategy avoids the problem of sorting through many models with exactly equal likelihoods. This approach still leaves a very large search space and is currently infeasible because it requires enumeration of the equivalence classes, which is unsolved. Because this approach relies on equivalence, negative rates must be allowed to guarantee that the highest likelihood is found. Nonequilibrium dynamical experiments might in principle be used to help choose among multiple models in the best-fitting equivalence class.

This strategy is very effective for rank 1 topologies because there is only one equivalence class for fixed \( N_o \) and \( N_c \). The identifiable rank 1 topologies all have the same number of links, \( N_o + N_c - 1 \), and thus the same BIC score. Moreover, they all have positive rates if DB is satisfied. Thus, there is no point in fitting to more than one rank 1 topology.

Whatever the search strategy, there is no point in searching nonidentifiable topologies, because they will always fit the data equally well in a continuum of ways and no better than the special case topologies that are identifiable. In general, it is not known which are the identifiable topologies, but we do know they can have no more than \( R(N_o + N_c - R) \) reactions, where \( R \) is the rank of the topology. Some of these topologies can be shown, using methods to be discussed elsewhere, to be nonidentifiable and may be omitted. Beyond rank 1, the space of such topologies is still exponentially large and contains many equivalences.

**Hierarchical Search**

We propose an alternative strategy. When fitting to a topology that contains the best topology as a subtopology, this fact should be indicated by very small rate constants for the links that should be removed. This observation suggests the following hierarchical search method:

1. Ascertain the rank, \( R \), by successive fits to MIR form for the various ranks. (\( N_o \) and \( N_c \) are assumed known from the dwell-time distributions.)
2. Fit to all topologies with \( R(N_o + N_c - R) \) links.
3. For fits that yield links with small rates, remove all combinations of those links and reevaluate the simplicity criterion (such as BIC) for the resulting topologies.

Again, for the sake of efficiency, one can omit remaining nonidentifiable topologies, to the extent they are known. We emphasize that the hierarchical search only helps to find simpler models; the fit will not improve beyond the MIR fits in step 1. In fact, we believe that the hierarchical strategy contains 200 million topologies for 47,677 rank 4 topologies with 16 links. In the latter case, there are >2.5 million connected topologies, so restricting to the 16-link topologies reduces the computational burden by a factor of >50. Unfortunately, we estimate that the search space for the hierarchical strategy contains 200 million topologies for \( N_r = N_c = 5 \) and 1 trillion for \( N_r = N_c = 6 \). The hierarchical strategy makes it practical to search up to \( N_r = N_c = 4 \).

The first step is always to estimate \( N_r \) and \( N_c \). We have shown that the rank \( R \) is also of fundamental importance. All three quantities can in principle be extracted from current traces without reference to any specific topology. Other biochemical data also might help to constrain these values. MIR form allows one to make use of the rank and is closer to BKU form than most proposals in the literature we examined. The number of topologies very close to MIR form, which will have few or no non-independent \( O-C \) links, is manageable. An effective strategy thus might be to start with MIR form and greedily move one link at a time, keeping the rank fixed.

The results presented here can be extended. In their current form they provide a number of useful tools for the identification of binary channel gating kinetics.

Note Added in Proof. It has come to our attention that Flomenbom et al. (32) give additional arguments and discussion relevant to the rank 1 equivalence theorem presented here.

J.E.P. thanks Andy Fraser for many useful conversations. We also thank Silvina Ponce Dawson, Bernardo Panto, Marty Golubitsky, Daniel Don-On Mak, and Werner Horsthemke for useful discussions. David Torney calculated the numbers of topologies in the discussion. We benefited from the QBSC course at State University of New York at Buffalo and from conversations with Fred Sachs, Feng Qin, and Anthony Auerbach and their students. We thank Francke Rinssland and Catherine Macken for suggestions on the manuscript. This work was supported by National Institutes of Health Grant 1R01GM65830-01 and Los Alamos National Laboratory-Laboratory-Directed Research and Development Grant XI8N under Department of Energy Contract W-7405-ENG-36.