Summary

Moerner [(2007) Proc Natl Acad Sci U S A 104, 12596–12602] points out the very rapid increase since 1990 of papers with “single molecule” in the title. Had he searched for “single channel,” or “single ion channel,” he would have found that the rapid increase occurred in the mid-1970s.

By the time single molecule work was starting in biochemistry, the single ion channel field had already developed the tools that were needed for analysis and interpretation of observations on single molecules.

Single ion channels have the great advantage that it is possible to obtain long electrophysiological recordings ($10^4$–$10^5$ transitions between open and shut), with good signal-to-noise ratio and good time resolution (down to about 20μs for channels with a “medium-sized” conductance, i.e., 40–50pS). For ion channels gated by a neurotransmitter, this can be done in the steady state or after fast concentration changes. In the case of the glycine channel, the information contained in steady-state experimental recordings has allowed up to 18 free parameters (rate constants) to be estimated [Burzomato et al. (2004) J Neurosci 24, 10924–10940; Colquhoun et al. (2003) J Physiol (Lond) 547, 699–728] for postulated mechanisms with 10 or more discrete states. Furthermore, this can be done with an exact allowance for missed events (events that are too short to be detected reliably because they are shorter than the experimental resolution).
Recently, there have been exciting advances in single molecule fluorescence methods. In the context of ligand-gated ion channels, such methods have enormous potential. It would be wonderful if it became possible to use fluorescence methods to measure, on a single molecule, the binding of the ligand (agonist) that activates the channel at the same time as measuring the opening and shutting of the channel. Such measurements could give an enormous amount of information about how the molecules work. Fluorescence records that resemble single-channel records have indeed been observed [e.g. Borisenko et al. (2003) Biophys J 84, 612–622; Sonnleitner et al. (2002) Proc Natl Acad Sci U S A 99, 12759–12764]. Up to now, these records are too short and have insufficient time resolution to achieve our aim, but at the present rate of progress our aim is not impossible.

In the simplest case, the fluorescence signal would change rapidly, in a stepwise fashion, when a ligand was bound, allowing measurement of the durations of times when a receptor was occupied and when it was vacant. In this case, no new theory would be needed because optimum fitting methods that can be used for any specified reaction mechanism already exist.

First, we shall review briefly some results obtained with single molecule fluorescence methods, with the aim of comparing them with single ion channel results. Next, we shall mention some of the stochastic theory that has been developed for the interpretation of single ion channel records and discuss how it might be applied to other sorts of single molecule observations. Finally, an example will be given of an analysis of an ion channel mechanism based on single molecule observations.

**Comparison of Fluorescence Methods with Single-Channel Recording**

Recordings of the current that flows through single ion channels show the alternation between the open and shut conformations of a single-channel molecule. A channel has several shut states, and usually several open states too. It would be very useful if methods could be devised to get information about transitions between states that are not distinguishable from one another in an electrical recording. To be useful, this information would need to have a resolution as good as that of the electrical record (10–40μs resolution).

In recent years, structural information has emerged about some ion channels, though this information is still incomplete in most cases. But crystallographic structures are static. What we need now is dynamic information on the conformational changes in the channel before it opens.
Optical methods are obvious candidates for this. Fluorescence techniques have been applied to ensembles of ion channels for about a decade (Mannuzzu et al., 1996; reviewed in Gandhi and Isacoff, 2005; and Zheng, 2006). These methods rely on having a fluorophore in a specific position on a recombinant channel. This can be done either by attaching a sulfhydryl-reactive reagent to a cysteine residue or by producing an appropriate fusion protein of the channel protein with a genetically encoded label such as one of the forms of green fluorescent protein (GFP). A rearrangement in the channel structure can affect the fluorophore properties if it changes the environment around the probe (making it more or less hydrophobic) or the accessibility of the probe to quenching by small molecules or ions that can be added to the medium. In addition, if a pair of fluorophores can be attached to the protein, changes in their distance may result and be measurable as Förster resonance energy transfer (FRET).

All these measurements can be done at the same time as the electrophysiological recording in whole cells or in excised patches. Nevertheless, this work is usually carried out on ensembles of hundreds of channels in order to ensure a reliable signal-to-noise ratio.

Other than single-channel recording, there are relatively few single molecule experiments on ion channels, possibly because of the difficulty of carrying out single molecule experiments in live cells or at least in a fairly intact biological membrane preparation (for a review see Sonnleitner and Isacoff, 2003).

**Single-Particle Tracking of Channels**

The biggest group of studies is the application of single-particle tracking to channels, in which single channels are observed for the purpose of studying their lateral diffusion in the cell membrane and their interaction with the cytoskeleton and specific proteins such as gephyrin.

Receptors can be tracked by imaging a single fluorophore attached to the receptor by an antibody (see for AMPA receptors Tardin et al., 2003). The limitation of using an organic dye as a reporter is that photobleaching severely limits the amount of data that can be acquired. Choquet and coworkers (Tardin et al., 2003) report obtaining single-receptor trajectories lasting up to 4 s (on average 244 ms; 33 Hz acquisition rate and 30 ms integration) with Cy5 or Alexa-647 dyes. Single-particle tracks lasting a few seconds were achieved for GABA_A receptors by labeling them with antibody-derivatized fluorescent nanospheres (0.1 μm in diameter) (Peran et al., 2001) and much longer times (tracks of 200 s) were achieved for glycine receptors by Triller and coworkers by using antibody-coated latex beads attached to an myc tag in the channel (Meier et al., 2001). Latex
beads are said to allow recording times of up to a couple of minutes before they stick to membranes (Dahan et al., 2003). However, they are large (0.5 μm in diameter), which reduces spatial resolution, may limit the probe access to the protein of interest, or indeed affect the very movements to be tracked by producing “drag.” A better compromise is that of using quantum dots, which are smaller (15–20 nm including the coating) and allow long recordings because of their resistance to photobleaching (reviewed in Bannai et al., 2006). However, they exhibit a property known as “blinking” in which their fluorescence output transiently blinks off and then on again. Although the blinking phenomenon can be exploited to confirm that a particular spot of light is produced by a single entity, the intermittent appearance and disappearance makes tracking and time series analysis more difficult. However, the excellent resistance of quantum dots to photobleaching means that recording periods are effectively limited simply by the cell survival during illumination. The time-lapse recording protocol described and discussed in Bannai et al. (2006) leads to a time resolution at 25 μs for a 5–10 nm point accuracy.

This group extended their previous work on glycine receptors in neurons in culture by using streptavidin-coated quantum dots attached to the receptor via a subunit-specific antibody and a biotinylated Fab fragment (Dahan et al., 2003). It was estimated that quantum dots allowed more than 20 min recording (vs. 5 s for Cy3-coupled antibody) and gave a signal-to-noise ratio an order of magnitude greater than Cy3 (Dahan et al., 2003). For extrasynaptic receptors, these measurements gave a diffusion coefficient value four times higher than the previous estimate from latex beads, suggesting that these had effectively “slowed” the receptor in the previous study. Whether this artifact always distorts results may depend both on the receptor and its effective diffusion rate and on the probe (see for instance Tardin et al., 2003). Even in these recent papers, the linker between the quantum dot and the receptor was still bulky, and further improvements may come from the adoption of quantum dots derivatized with peptide receptor ligands such as those that have been developed for G-protein–linked receptors (Zhou et al., 2007). For a discussion of the analysis of such results see Bouzigues and Dahan (2007) and Ghosh and Wirth (2007). Techniques for single-particle tracking have been used for a variety of channels, including AMPA, NMDA, GABA\textsubscript{A}, and potassium channels (see for instance Borgdorff and Choquet, 2002; Bouzigues and Dahan, 2007; Groc et al., 2006; O’Connell et al., 2006). For more on this subject, see also Chapter 9, this volume.

**Stoichiometry**

While photobleaching limits the duration of data collection from a fluorophore molecule, it has been turned to good use for the purpose of counting subunits in channels by
Ulbrich and Isacoff (2007). This is an elegant technique, which involves the imaging by total internal reflection fluorescence (TIRF) microscopy of recombinant channel molecules expressed in *Xenopus* oocytes at low expression density. The fluorescence signal comes from GFP tags inserted in the channel subunits, and the stoichiometry of a single channel is assessed by counting individual steps in the emission as the fluorophore is photobleached. Direct counting is not entirely straightforward because approximately 20% of GFP tags are not fluorescent, maybe because of defects in folding during assembly. Furthermore, two bleaching steps may appear as one in the record because of temporal proximity (the authors estimate that the dwell times for each level of fluorescence observed with their protocol imply that 10% of steps would be missed). Even so, the results of the technique appear to be robust and reliable if proper control experiments are carried out.

**Imaging Ion Fluxes Through Single Ion Channels**

Fluorescent calcium ion indicators have allowed the visualization of localized calcium ion transients (sparks) for a long time (Cheng et al., 1993), but it is only recently that technical refinements have allowed restricting the volume to be probed to that immediately adjacent to the plasma membrane. It is in this narrow volume that changes in calcium concentrations can be expected to follow closely the activity of single ion channels (Shuai and Parker, 2005).

An excellent example of what is now possible is given by the work of Demuro and Parker (2003, 2005) on voltage-gated calcium channels and muscle nicotinic acetylcholine receptors. The authors have used TIRF microscopy to restrict the excitation of fluo-4 dextran loaded in a *Xenopus* oocyte to the 100nm thickness of the evanescent wave and obtained simultaneous recordings from many (100 or more) individual channels. Muscle nicotinic channels appear to be randomly distributed in the oocyte membrane and showed little lateral movement over tens of seconds. Only stage VI oocytes were used, as these have lost most of the membrane microvilli. It is not clear whether all the functional channels are imaged or whether the remaining microvilli restrict the proportion of imaged channels to those in favorable locations.

A highly sensitive CCD camera made a frame rate of 500 s⁻¹ possible with good signal-to-noise ratio, and the factors that limit the time resolution on these records become the affinity of the fluorescent dye and the volume of the intracellular space sampled in the process. Shuai and Parker (2005) estimated that 1–2 ms resolution is possible with the best “compromise” choices, in particular with a fluorescent dye affinity for calcium
ions of between 1 and 5 μM (a higher affinity would slow equilibration and a lower one would reduce sensitivity). While this is much less than the resolution achievable with electrophysiological recording, the “optical patch” technique allows the activity of a particular single-channel molecule to be followed in time. This is not possible in normal patch clamping (unless you can be sure that the patch contains only one channel, which is something that cannot be easily achieved for ligand-gated channels). It is also remarkable that the channels in the optical patch remained accessible to drugs applied to the bulk bath solution, despite the fact that the TIRF conditions require the oocyte membrane to be resting on the coverslip of the microscope. Irrespective of the time resolution, this method may also allow counting of the number of functional channels in an area (assuming no channel is missed because of its position in a membrane invagination outside the optical “slice” that is excited). It would be very useful to be able to count the number of functional channels in a patch from which channels are being recorded by patch clamp, but the geometry of a normal patch (i.e., a dome on the tip of a glass pipette) makes TIRF imaging impossible. The only possibility would be if the recent development of planar patching devices (Klemic et al., 2002) could be extended to make it feasible to obtain electrophysiological recordings from planar patches held close (≤100 nm) to the refractive interface needed for TIRF.

**Detecting Conformational Changes in Single-Channel Molecules**

Proof of principle that this can be achieved was provided in Shaker potassium channels by Sonnleitner et al. (2002), who labeled the S4 voltage-sensing domain of these channels with tetramethylrhodamine (at very low concentration, to ensure that the optical patch would have a low density of fluorophores and that each channel would bear one dye molecule at most on its four S4 domains). The position of the labeled residue was chosen among those that gave a large fluorescence change upon channel activation in ensemble recordings. Changes in the fluorophore emission in response to voltage steps applied by two-electrode voltage clamp were measured by TIRF microscopy of the oocyte membrane. Once again, TIRF techniques were essential to reduce background fluorescence and this required careful preparation of the oocyte, including removal of both the vitelline membrane and the extracellular matrix by enzyme treatment. It is likely that the imaging monitors only channels located at the tip of microvilli in the oocyte membrane. Because several channels were imaged at one time (by a CCD camera with a 25 ms frame time), the time resolution was less than the theoretical maximum of the technique. The same authors (Sonnleitner and Isacoff, 2003) estimate that this could be 10 μs, if only one channel is monitored at a time with an avalanche photodiode and if the fluorophore is excited at a near-saturating level. If that were possible, the resolution would become comparable with
that of electrical recording. However, photobleaching means that there would be a trade-off between temporal resolution and the average duration of the optical recording. If a fluorophore molecule emits $10^5$–$10^6$ photons before irreversible damage, and 100 photons per data point are needed to make the measurement signal-to-noise ratio acceptable, only 1000–10000 data points can be obtained (Ha, 2001) (Figures 8.1 and 8.2).

**Simultaneous Electrophysiological and Optical Recording from Single Channels**

The suggestion that single-channel recording and single-pair FRET could be advantageously combined was made as early as 1999 (Weiss, 1999), 3 years after the first demonstration of single-pair FRET (Ha et al., 1996), but to our knowledge there are only two cases in which this “holy grail” has been achieved. Both studies monitored the single-channel activity of gramicidin in artificial bilayers and have used single-pair FRET to determine the dimerization of gramicidin (Borisenko et al., 2003; Harms et al., 2003).

![Figure 8.1](image.png)

**Figure 8.1:** An experimental setup for total internal fluorescence microscopy (TIRF) on oocytes. The line drawing shows the arrangement of a Nd:YAG laser for fluorophore excitation through a single-mode optical fiber for spatial filtering and expansion, a shutter (S), a polarizer (P), and a dichroic mirror (DM). Detection is with an image-intensified cooled charge-coupled device (ICCD), via the appropriate filters (F) and a mirror (M). Channels are activated by voltage steps applied by a two-electrode voltage clamp. Reproduced with permission from Sonnleitner et al. (2002), where further details of the method can be found.
Figure 8.2: Fractional fluorescence changes upon voltage-gated conformational change in single potassium channel molecules. Oocyte-expressed Shaker-type potassium channels were labeled at a specific cysteine residue in their voltage-sensing domain with tetramethylrhodamine-5-maleimide. The frames in the first row are averages \((n = 2, 30 \times 30\) pixel, corresponding to \(6.75 \mu m \times 6.75 \mu m\)) taken immediately before (A) or immediately after (B) the oocyte is repolarized from +20 to −100 mV. C is the difference between B and A. The intensity of fluorescence is measured from this sort of image at the circles over a period of 10 s, and examples of different trajectories are shown in the three rows below (D–J), together with the steps in the command voltage (from −100 to +20 mV, 1 s pulses at each voltage). Some of the trajectories (numbers) were obtained from the optical patch in A–C. Some spots (see traces, D and G) are bright throughout the voltage perturbation and are likely to be nonspecifically
The experimental details are somewhat different: Borisenko et al. (2003) added to planar bilayers Cy3 gramicidin (donor) and a conductance mutant form of Cy5 gramicidin (acceptor) in order to be able to distinguish in the electrophysiological recordings the heterodimers (which are the only channels that have a donor and an acceptor and therefore could give rise to FRET) from homodimer channels (i.e., containing two acceptor or two donor molecules).

In contrast, Harms et al. (2003) allowed only heterodimers to form, as the gramicidin tagged with the donor (tetramethylrhodamine) and that tagged with the acceptor dye (Cy5) were added from different sides of the bilayer, which was formed on the tip of a patch pipette. In both studies, images were taken with a 5 or 6 ms exposure, every 7–1000 ms. The temporal resolution is therefore far slower than can be attained with electrical measurements. Both groups report that some openings failed to be associated with detectable FRET events. This could be for a variety of reasons. For example, it could result from blinking or photobleaching of one of the dyes in the pair, incomplete imaging of the membrane (especially at the edges of the field), and the possibility that in some dimerized channels or in some states of the open channels the angle between the fluorophores is inefficient for FRET. Conversely, FRET events were observed in the absence of openings. Again it is impossible to be sure of the explanation for this, which could be that dimers can form in parts of the image where there is no bilayer as such, or that the dimers do not conduct because they represent intermediate conformations that precede opening.

This brief survey shows that the potential of fluorescence methods for contributing to improvement of channel kinetic measurements has yet to be fulfilled, but also that there is hope that it soon might be. This topic is discussed further in Chapter 9, this volume. Next, we shall consider the extent to which the existing single molecule stochastic theory, and existing methods for extracting information from experimental data, may be useful for optical measurements from single molecules.

bound rhodamine molecules. Other dim spots (white circle, D) also represent background signal. Channels labeled with a fluorophore in this position responded to depolarization with a decrease in fluorescence and are the bright spots in the difference image, C. After a short time, voltage pulses cease to elicit fluorescence changes, probably because of bleaching (arrows and asterisks). Recording at 20 frames per second, high-intensity illumination: 1.2 mW per 20 μm diameter spot. Reproduced with permission from Sonnleitner et al. (2002)
When interpreting the interpretation of open an have 0, 1, or 2 molecules. An example is shown in Figure 8.3. For example, the states may differ because they have 0, 1, or more ligand molecules bound, or because the channel is in the open or shut conformation. The observed electrophysiological record distinguishes only between the open and the shut states.

In the interpretation of macroscopic recordings, that is, from populations of channels, it is supposed that the transitions between these discrete states obey the law of mass action. When dealing with a single molecule, the usual kinetic analyses have to be recast in the form of probabilities. This was done by Colquhoun and Hawkes (1977, 1982). In order to do this, it is supposed that the transitions between states are a Markov process¹ (this implies the law of mass action, though the converse is not necessarily true).

¹ A Markov process is a random (stochastic) process that has no “memory,” in the sense that the future evolution of the system depends only on its present state, not its past history. Put another way, events that occur in nonoverlapping time intervals are independent.
The only way to write the theory for anything but the very simplest mechanisms is to use matrix notation.\(^2\) This has the great advantage that expressions look very simple, while at the same time being entirely general (e.g., Colquhoun and Hawkes, 1995a,b).

At the heart of every application of this sort of analysis lies a table that contains all of the rate constants for transitions between all possible pairs of states. The table, usually called the \(Q\) matrix, defines the mechanism completely.

**The \(Q\) Matrix Defines the Mechanism**

Consider, as an example, the mechanism shown in Figure 8.3A. The diagram shows the simplest mechanism for a (muscle-type) nicotinic acetylcholine receptor. The agonist (e.g., acetylcholine) is denoted \(A\), and the receptor (which has two binding sites for the agonist) is denoted \(R\) (when shut) or \(R^*\) (when open). The mechanism has five discrete states, numbered 1–5, and the names of the rate constants for transitions between these states are written on the arrows. These rate constants are written down in a table (the \(Q\) matrix, Figure 8.3B), such that the entry in the \(i\)th row and the \(j\)th column is the transition rate from state \(i\) to state \(j\). For example, the transition rate for opening of the doubly liganded receptor, \(\beta_2\), is the transition from state 3 to state 2, and so appears in the third row and second column. Also notice that zero entries correspond to states that are not connected (for instance, states 5 and 1, see entries on fifth row and first column). The \(Q\) matrix defines the mechanism entirely. Incidentally, the transition rate from state 1 to state 1 means nothing; the elements on the diagonal are defined (simply for convenience) so that the rows add up to zero.

**Macroscopic Experiments**

The formulation is incredibly simple if we are recording from many molecules. We want to know how the fraction of molecules in each state varies with time after a perturbation (e.g., a concentration jump or a jump in membrane potential), and what equilibrium occupancies are eventually achieved. Call the occupancies of the five states \(p_1(t), p_2(t), \ldots, p_5(t)\), at a time \(t\), and write these in a table (a row vector) denoted \(\mathbf{p}(t)\). Thus

\[
\mathbf{p}(t) = \begin{bmatrix} p_1(t) & p_2(t) & p_3(t) & p_4(t) & p_5(t) \end{bmatrix}
\]  

(8.1)

---

\(^2\) Most biologists are not familiar with matrix notation, but we have found that it is quite feasible to get a working knowledge of the subject in a 5-day graduate school course, for example.
Whole books have been written about macroscopic kinetics, but they all boil down to special cases of the quite general solution to the problem

\[ p(t) = p(0) \exp(Qt) \]  

(8.2)

where \( p(0) \) represents the occupancies at zero time (the time when the perturbation is applied). This solution holds as long as \( Q \) is constant, that is, the transition rates are all constants. In practice, this simple expression is most conveniently separated into \( k - 1 \) exponential components (where \( k \) is the number of states in the mechanism), plus an equilibrium component. The equilibrium occupancies, \( p(\infty) \), are defined by

\[ p(\infty)Q = 0 \]  

(8.3)

In practice, it always turns out, except in the simplest cases, that macroscopic recordings do not contain enough information to define all the transition rates in the mechanism. Single molecule measurements contain more information and may be able to define the mechanism better.

From macroscopic kinetics, we now move to the more interesting topic of single molecules.

**Single Molecule Experiments**

Figure 8.4 shows part of an experimental single-channel recording.

The baseline of the record (top) represents the condition when the receptor is in one or other of the shut states. There is no way to tell which of the shut states it is in at any particular moment (there are three shut states in Figure 8.3A, for example). For that reason the process is described as an *aggregated Markov process*, or, sometimes, as a hidden Markov process, though the latter term is used in more than one sense, so is less desirable.

Similarly, the upward deflections show times when the channel is open, that is, in one or other of the open states (again we cannot tell which).

It is characteristic of a Markov process that the distribution (probability density function) for the length of time spent in any individual state is a simple exponential distribution

\[ f(t) = \lambda \exp(-\lambda t) \]  

(8.4)

where \( \tau = 1/\lambda \) is the mean lifetime of the state in question. But for an aggregated process it gets slightly more complicated because what we observe (the duration of an opening or a shutting) represents the sojourn of the channel in a *set* of states (either open states
or shut states) and this dwell time can be made up by many individual visits to different states in that set.

Define a set A that contains all the open states (states 1 and 2 in Figure 8.3) and a set F that contains all the shut states (states 3, 4, and 5 in the example). The distribution of open times can then be written, quite generally, in the very simple form (Colquhoun and Hawkes, 1982)

\[ f(t) = \varphi_A \exp(Q_{AA}t)(-Q_{AA})u_A \]

In this expression, \( Q_{AA} \) is the subsection of the \( Q \) matrix that refers to transitions between open states. For example, for the mechanism in Figure 8.3A, the partition of \( Q \) into open and shut states is marked in Figure 8.3B and

---

\[ ^3 \text{The sets of open and closed states might, more obviously, be denoted O and C, but the notation A and F is widely used, for historical reasons, and emphasizes that the results are valid for any set of states that are distinguishable in the experimental record.} \]
\[ Q_{AA} = \begin{bmatrix} -(\alpha_1 + k_+^s[A]) & k_+^s[A] \\ 2k_-^s & -(\alpha_2 + 2k_-^s) \end{bmatrix} \] (8.6)

The only other things that are needed are an initial vector, \( \varphi_A \), that contains the probabilities that an opening starts in each of the open states (this can also be calculated from the \( Q \) matrix, see Colquhoun and Hawkes, 1982), and a vector \( u_A \) with elements all 1. This expression is directly analogous with the simple exponential distribution, Eq. (8.4), except that \(-Q_{AA}\) is in place of \( \lambda \), but it is entirely general, for any mechanism, however many states it may have.

The distribution of shut times is exactly analogous: all you have to do is to interchange A and F in Eq. (8.5).

Exactly the same expressions can be used for any subset of states that is distinguishable in the data. Imagine, for example, that a fluorescence record could be obtained that showed step changes when a binding site became occupied. Then, set A could be defined as the set of occupied states, and the distribution of the durations of occupancies could be calculated from Eq. (8.5).

**Fitting Data**

The problem in real life is to decide what mechanism describes your experimental results. In the early days of single-channel analysis, this was done by counting the number of components in the various types of dwell-time distributions and patching together the information into a plausible mechanism. This method produced the first dissection of agonist affinity and efficacy in a ligand-gated ion channel (Colquhoun and Sakmann, 1981, 1985), but it has several limitations. For example, analysis of distributions does not make use of all the information in the records; the appropriate number of components in the distributions can be ambiguous, estimating the underlying rate constants from observed time constants is indirect and difficult, and a proper correction for missed events is not possible until a mechanism is specified.

The modern strategy is to start by postulating a mechanism, obtain the values of the rate constants that fit your results best, and then see if the fit is adequate.

Several methods have been proposed for fitting, but only two have been used for real experiments. They are both based on maximizing the likelihood of the observed sequence of open and shut times (or durations of sojourns in A and not-A in general). The observed record must first be idealized (Colquhoun and Sigworth, 1995) to produce a list of open and shut times. The calculation of the likelihood of this observed sequence is not unambiguous.
Two programs are freely available for doing this: QUB (http://www.qub.buffalo.edu/) and HJCFIT. They use somewhat different methods for doing the calculations, and a comparison of the pros and cons for each method is available at http://www.ucl.ac.uk/Pharmacology/dcpr95.html#hjcfit. The only method for which the performance has been properly tested is HJCFIT (Burzomato et al., 2004; Colquhoun et al., 2003).

Other methods of fitting experimental data have been suggested for electrophysiological records from ion channels. In particular, hidden Markov methods (HMM) have been proposed (Qin et al., 2000; Venkataramanan et al., 1998a,b, 2000) for analysis of single ion channel data from electrophysiological recording. It has also been suggested as a method for fitting FRET records (McKinney et al., 2006). Whatever theoretical virtues HMM may have, they are far too slow for routine use on real data, as the probability of being in each state has to be calculated for every individual sample point in the raw data (3 × 10^7 points in a 10min record at 50kHz). In any case, HMM have never been tested by analysis of repeated simulated experiments, so it is not known whether they have any advantage or not.

The Number of Channels in a Patch

More information could be obtained from single-channel recordings if we knew how many channels were present. Occasionally that can be estimated, but usually it cannot. This has led to development of methods for the analysis of single-channel records that work even when the number of channels is not known. Part of the analysis consists of defining sequences of events in the experimental record that are likely to come from the same molecule. For instance, ligand-gated ion channels of the nicotinic or glutamate superfamily at low concentrations of agonists open in bursts, where a burst is essentially the events between the binding of the first molecule of ligand and the unbinding of the last molecule. Bursts of openings are separated by long shut times, and it is impossible to be sure that one burst originates from the same individual channel as the burst that preceded it. This means that the information contained in the shut times between the bursts cannot be used for fitting (because if there is more than one channel present, these shut times will be shorter than expected from the mechanism), and the likelihood must be calculated separately for each burst. A method has been devised for recovering some of the information lost in this way, by using “CHS vectors” for calculation of the likelihood (Colquhoun et al., 1996), and this has proved to be useful in practice (Colquhoun et al., 2003).

It would be still better to know the number of channels. In principle, one might imagine that a combination of fluorescence and patch-clamp methods might be able to tell one how many channels were present, but this has not been achieved yet.
The HJCFIT Method

The name of this program is derived from the papers by Hawkes, Jalali, and Colquhoun (Hawkes et al., 1990, 1992) that provided the exact solution to the missed events problem. In practice, it is not possible to detect very short events (in our case, very short openings or shuttings). In our single-channel recordings, the shortest resolvable event is usually 20–40 μs, although with suboptimal methods of idealization (or with channels with small conductances) it may be much longer. Since shut time components with mean durations in the 10–20 μs range seem to be common in many sorts of channel, many events may be missed and accurate allowance for this is essential. For example, the more the shuttings that are missed, the longer the openings will appear to be. What we need is the distribution of apparent open and shut times, as extended by missed shuttings and openings, respectively. Our way of calculating these gives what we refer to as HJC distributions. These describe what is actually observed under these circumstances, rather than the simple results such as Eq. (8.5), which would apply only with perfect resolution. The correction can be quite big. For example, the mean apparent open time may be about 20-fold longer than the true open time at high glycine concentrations (Burzomato et al., 2004; see Figure 8.6). If events are not missed, the form of the distribution of dwell times is a mixture of exponentials under our conditions, but HJC distributions have not got this form for the shortest events (below about three times the resolution).

Despite these complications, the practice of fitting a mechanism to the observations is reasonably straightforward. Records are obtained at a range of agonist concentrations and idealized, namely, turned into lists of openings and shuttings, including their duration. At this stage, the effective time resolution of the record is estimated. A preliminary analysis is carried out in order to identify which stretches of the record can be assumed to come from the same channel. At high agonist concentrations, these can be very large groups of openings (clusters) identified by the lack of double openings4 despite relatively high open probability. At low agonist concentrations, only short sections of the record (bursts) can be assumed to come from the same individual channel.

Everything is now ready to be entered into the HJCFIT program. The program calculates the likelihood of the observations (typically several sets obtained at different agonist concentrations) from the postulated mechanism, the resolution, and a set of guesses for the rate constants (i.e., the elements of the $Q$ matrix). The rate constant values are then changed until that likelihood is maximized. The QUB program also maximizes likelihood,

---

4 Indicated by double the normal channel current.
but it calculates the likelihood in a way somewhat different from ours. In order to judge how well the mechanism fits the data, the original observations are then displayed as dwell-time distributions, conditional distributions, and $P_{\text{open}}$ (probability of being open), and predictions of the fit (HJC distributions) are superimposed on the observations. The plausibility of the mechanism fit is judged by the extent to which these predictions superimpose on the data.

In principle, the calculation of the likelihood is simple. We define a matrix, $G_{\text{AF}}(t)$, which contains (roughly speaking) the probability densities for staying within the subset of states A (e.g., open) for a time $t$ and then leaving A for a state not in A (in subset F). Colquhoun and Hawkes (1982) show that this is given by

$$G_{\text{AF}}(t) = \exp(Q_{\text{AA}}t)Q_{\text{AF}}$$

(8.7)

where $Q_{\text{AF}}$ is simply the top right subsection of the $Q$ matrix, as illustrated in Figure 8.3B. $G_{\text{FA}}(t)$ is defined simply by interchanging A and F in Eq. (8.7).

The likelihood, $l$, of a whole sequence of observed (apparent) open and shut times can now be calculated, as described by Colquhoun et al. (1996).

$$l = \varphi_A G_{\text{AF}}(t_{o1}) G_{\text{FA}}(t_{s1}) G_{\text{AF}}(t_{o2}) G_{\text{FA}}(t_{s2}) G_{\text{AF}}(t_{o3}) \ldots u_F$$

(8.8)

where $t_{o1}, t_{o2}, \ldots$ represent the first, second apparent open time, and $t_{s1}, t_{s2}, \ldots$ first, second apparent shut time, etc. Note that openings and shottings occur in this expression in the order in which they are observed. Thus, $\varphi_A$ is a $1 \times k_A$ row vector giving the probabilities that the first opening starts in each of the open states. The first two factors, $\varphi_A G_{\text{AF}}(t_{o1})$, form a $1 \times k_F$ row vector the elements of which give the probabilities that the first shut time, $t_{s1}$, starts in each of the shut states, at the end of an open time of length $t_{o1}$. Then the first three terms, $\varphi_A G_{\text{AF}}(t_{o1}) G_{\text{FA}}(t_{s1})$, form a $1 \times k_A$ row vector the elements of which give the probabilities that the next open time, $t_{o2}$, starts in each of the open states, after an opening of length $t_{o1}$ and a shutting of length $t_{s1}$. And so on up to the end of the observations, where the elements of the last row vector are summed over shut states to give the (scalar) value of the likelihood by (post)multiplying the row vector by $u_F$, the unit column vector (size $k_F \times 1$). The process of building up the product in Eq. (8.8) gives, at each stage, the joint density of the time intervals recorded thus far, and it can be regarded as progressive calculation of a vector that specifies probabilities for which state the next interval starts in, which will depend on the durations of those intervals. This process uses all the information in the record about correlations between intervals.
In real life, the calculation of $G_{AF}(t)$ and $G_{FA}(t)$ is more complicated than using the simple result in Eq. (8.7), because their definition must be changed to allow for missed events, as described by Colquhoun et al. (1996).

Fitting of 9465 resolved intervals with 9 free parameters takes less than 30 s on a standard PC. The fit produces estimates of all the rate constants, that is, of the $Q$ matrix. It also produces estimates of their errors and their correlations (their covariance matrix). At this stage the predictions of the fit are compared with the data. In this method there is no need to fit separately distributions of apparent open times, shut times, burst lengths, etc. (and so no need to agonize about how many exponentials to fit). Rather, all the information in the record is taken into account simultaneously to get estimates of the $Q$ matrix, and the predictions calculated from these estimates are superimposed on the data to see how successful the fit was. Examples are shown in the next section.

**Some Recent Results from Single-Channel Recording**

The high resolution of single-channel recording methods has allowed quite complicated reaction mechanisms, with up to 18 free parameters, to be estimated from experimental data, with reasonable precision (Burzomato et al., 2004). That is far more than can be defined by any macroscopic method, for which it is rare for even five free parameters to be estimatable. Furthermore, rate constants up to 130 000 s$^{-1}$ can be found reliably, and that is much faster than can be found by other estimation methods. In particular, the rival QUB method does not seem to be capable of getting estimates of opening rate constants from low-concentration data.

The fitting process will be illustrated by an example from a study designed to cast light on the mechanism of the activation of the heteromeric glycine receptor by the neurotransmitter glycine and partial agonists.

After fitting over 30 different putative mechanisms, the choice was reduced to 2 (shown in Figure 8.5A and B). The results could be fitted by mechanisms with three binding sites (but not by those with only two). That is not surprising in view of the reported $\alpha 1_3\beta 2$ subunit composition. With three binding sites, the minimum mechanism would have four shut states (with 0, 1, 2, or 3 glycine molecules bound) and three open states (the lower two rows in Figure 8.5A). But it was found that four shut states are not enough to give a good fit over the whole concentration range (analogous results have been found with the nicotinic receptor and the BK channel).

The simplest way to add extra shut states to the mechanism is to have just one connection between each additional closed state and the original scheme, as this puts the least
Figure 8.5: Activation mechanisms for the heteromeric glycine channel. Reproduced with permission from Burzomato et al. (2004), see text for a detailed description.

Although a good fit was obtained (Burzomato et al., 2004), the approach of simply adding unconstrained shut states is somewhat unsatisfactory, for several reasons.

First, the additional shut states are added arbitrarily to get a fit. Do these states correspond with physical reality? Although they are labeled D to suggest that they are desensitized states, they are all short-lived and do not account for the desensitization seen with macroscopic methods.
Second, for the glycine receptor we found that a good fit with the scheme in Figure 8.5A is achieved only if binding sites are allowed to interact to produce a strong increase in affinity as more ligand molecules bind (what is commonly, but unhelpfully, called “cooperativity of binding”). Thus, the first binding has low affinity \( K_1 = 14\,000 \mu M \), but once one site is occupied, the next binding has much higher affinity \( K_2 = 200\,\mu M \), and when two are already occupied, the third has even higher affinity \( K_3 = 10\,\mu M \).

The problem with this finding is that it is not clear how this interaction occurs, given the relatively long distance between the binding sites in the channel. How can one site detect when another is occupied, if the channel is supposed not to have undergone any major change of conformation?

Both of these objections to the mechanism in Figure 8.5A are removed if the three “extra” shut states are inserted between the resting state and the open state to give the “flip” mechanism in Figure 8.5B. This mechanism fits the data as well as that in Figure 8.5A, despite having four fewer free parameters. Fitting mechanisms of this sort is useful only if the states that are postulated have real physical existence. The flip mechanism has the advantage that it makes a plausible postulate about the physical nature of the channel activation process. The agonist binding site is some distance from the channel gate, and it has been suggested that there must be intermediate states between the initial binding and the final opening of the channel (e.g., Colquhoun, 2005; Grosman et al., 2000). The flip mechanism postulates that such an intermediate state lasts sufficiently long to be detected in the observations. The extra flip conformation states represent a (concerted) change in conformation that occurs after the agonist binds but before the channel opens.

One plausible speculation would be that this transition represents the “domain closure” seen in crystal structures of the extracellular domain of glutamate channels (Jin et al., 2003). In the “flip” mechanism, the apparent increase in affinity with increase in ligand formation is explained in the same way as was originally postulated for hemoglobin (Wyman and Allen, 1951). The binding sites behave independently and do not interact, but can exist in two conformations, resting (low affinity) and “flipped” (higher affinity).

Within each conformation, the affinity of each binding step is the same. For example, for the glycine receptor we estimated that the affinity (as measured by the dissociation equilibrium constant) is \( 520\,\mu M \) for the first, second, and third binding to the resting conformation. Similarly, it is also postulated that the affinity is the same for each successive binding to the flipped conformation (F). But the affinity for the flipped conformation \( (8\,\mu M) \) is 65 times greater than for the resting conformation, so binding favors the higher affinity F state and hence activation of the receptor.
From the point of view of affinity and efficacy (Colquhoun, 1998), we see that there are now two different affinities, but only one of them tells us about the resting state of the receptor, so \( K = 520 \mu M \) is the “affinity” in the sense that Stephenson (1956) originally intended (but failed to measure). The efficacy now involves two different steps, flipping and opening. Although the gating constant still increases with the number of agonist molecules bound, it does not increase as much as the flipping constant (65-fold increase for each ligand bound). So, according to this interpretation, flipping (while shut) is more important in determining efficacy than the opening reaction itself.

**How the Fit is Judged**

Figure 8.6 shows how the results of a mechanism fit can be displayed. The data fitted were obtained at four glycine concentrations and fitted with the “flip” mechanism in panel A. Panels B and C show different ways of plotting the recorded data \( P_{\text{open}}(P_{\text{dose}}) \) dose–response curve, open and shut time distributions, and conditional mean open time plots. The last of these is a measure of the (negative) correlation that is observed between the length of an opening and the length of the adjacent shut time. In each case, the actual data are shown together with the predictions for that distribution, calculated from the rate constant values obtained from fitting the mechanism to the idealized records. Note that, for example, the distribution of apparent open times has not been fitted directly, but the solid line that has been superimposed on the observed histogram is the HJC distribution of apparent open times that has been calculated from the optimum values of the rate constants found in the fit, which was carried out on a set of four idealized records together with the mechanism and the resolution (30\( \mu s \) in this case).

In Figure 8.6, the dashed lines show the predictions of the fit for what would be seen with perfect resolution, as calculated from the fitted rate constants by the simple results such as Eq. (8.5). The extent to which the ideal (dashed line) distribution lies to the left of the HJC distribution (solid line) is a measure of the distorting effect of missed short events. This is particularly obvious for open times in Figure 8.6B. The apparent open time is roughly 20-fold more than the true open time at 1000\( \mu M \) glycine, because of the large number of missed brief shuttings. This emphasizes the importance of making accurate corrections for missed events. The conditional mean open time plots (bottom row) show the observations, which are the mean duration of those openings that are adjacent to shut times in a specified range, as filled circles joined by a solid line, and the predicted correlations as open symbols and dashed lines.

The purpose of these displays is to make it visible to the eye how well or how badly the mechanism describes the data.
Figure 8.6: The “flip” mechanism provides an accurate description of the single-channel activity of the glycine heteromeric channel. Reproduced with permission from Burzomato et al. (2004), see text for a detailed description
Implications of the “Flip” Mechanism

In contrast to the heteromeric receptor that has just been described, the homomeric glycine receptor, when interpreted in terms of the flip mechanism, shows a smaller affinity difference between the resting and flipped conformations of the receptor (6-fold rather than 65-fold; Burzomato et al., 2004). This explains the smaller “apparent cooperativity” of the homomeric receptor. It also shows how unhelpful the term “cooperativity” is when discussing mechanisms.

A mutant glycine receptor that causes an inherited defect in the effectiveness of glycnergic synapses in mice, α1A52Sβ, shows hardly any difference in affinity between resting and flipped conformations (Plested et al., 2007). This means that on mutant channels the neurotransmitter is not as effective in producing the conformational changes that precede opening and provides a new way of interpreting the effects of mutations.

It has been supposed, ever since the work of del Castillo and Katz (1957), that partial agonists produce a small maximum response because of inefficiency of the gating step (the major conformation change that accompanies the opening and shutting of the channel). We now have evidence, for both nicotinic and glycine receptors, that the actual open-shut step is very similar for both full and partial agonists (Lape et al., 2008). What limits the maximum response for partial agonists is not gating but flipping.

According to this interpretation, it seems likely that the root of partial agonism lies before the gating step, and perhaps, therefore, nearer to the extracellular binding site region. This has interesting implications for structural studies. There seems to be little hope of relating structure and function if we have not got a description of function that describes the actual physical events that occur during channel activation.

Conclusions

The example described above shows how much has been achieved in understanding how ion channels work in the 30 years that have elapsed since the current signal produced by a single ion channel molecule was first recorded in a physiological membrane (Neher and Sakmann, 1976). Single molecule currents had been observed even earlier in lipid bilayers with gramicidin (Hladky and Haydon, 1970, 1972) and “excitability inducing material” (Latorre et al., 1972).

Other single molecule techniques are increasingly being applied to ion channels, which raises the hope that they soon will contribute information to complement and enrich that obtained from electrophysiology. It would, for example, be very useful to have a way to
count the number of functional channels in the patch from which currents are recorded. More ambitiously, it would be very useful to have additional time-resolved information that would allow us to “dissaggregate,” at least in part, the different states in the sets of our Markov mechanisms. In the future, we expect that the single molecule optical methods described earlier will enable us to address some of these critical questions. An ability to make both electrical and optical measurements simultaneously would certainly advance our basic understanding. For any type of channel, this means finding a way to detect and report conformational changes short of opening at a time resolution similar to that achieved by electrical recordings. In addition to that, for ligand-gated ion channels it would be helpful to be able to detect whether one or more binding sites are occupied. The theoretical basis for analysis of single ion channel signals has been developed since 1977, and has now reached a state where mechanisms of any complexity can be fitted to data. Existing theory could cope with other single molecule measurements of the sort described above with little or no change.

References


Chapter 8


