Measuring Kinetics of Complex Single Ion Channel Data Using Mean-Variance Histograms

Joseph B. Patlak
Department of Physiology and Biophysics, University of Vermont, Medical Research Facility, Colchester, Vermont 05446 USA

ABSTRACT The measurement of single ion channel kinetics is difficult when those channels exhibit subconductance events. When the kinetics are fast, and when the current magnitudes are small, as is the case for Na^+, Ca^{2+}, and some K^+ channels, these difficulties can lead to serious errors in the estimation of channel kinetics. I present here a method, based on the construction and analysis of mean-variance histograms, that can overcome these problems. A mean-variance histogram is constructed by calculating the mean current and the current variance within a brief "window" (a set of N consecutive data samples) superimposed on the digitized raw channel data. Systematic movement of this window over the data produces large numbers of mean-variance pairs which can be assembled into a two-dimensional histogram. Defined current levels (open, closed, or sublevel) appear in such plots as low variance regions. The total number of events in such low variance regions is estimated by curve fitting and plotted as a function of window width. This function decreases with the same time constants as the original dwell time probability distribution for each of the regions. The method can therefore be used: 1) to present a qualitative summary of the single channel data from which the signal-to-noise ratio, open channel noise, steadiness of the baseline, and number of conductance levels can be quickly determined; 2) to quantify the dwell time distribution in each of the levels exhibited.

In this paper I present the analysis of a Na^+ channel recording that had a number of complexities. The signal-to-noise rato was only about 8 for the main open state, open channel noise, and fast flickers to other states were present, as were a substantial number of subconductance states. "Standard" half-amplitude threshold analysis of these data produce open and closed time histograms that were well fitted by the sum of two exponentials, but with apparently erroneous time constants, whereas the mean-variance histogram technique provided a more credible analysis of the open, closed, and subconductance times for the patch. I also show that the method produces accurate results on simulated data in a wide variety of conditions, whereas the half-amplitude method, when applied to complex simulated data shows the same errors as were apparent in the real data. The utility and the limitations of this new method are discussed.

INTRODUCTION

Recording the currents through single ion channels has become a firmly established method for determining both the conductance and kinetic properties of the channels. Even though many aspects of such analyses can be automated so as to decrease the tedium of sorting through thousands of channel openings in each experimental condition, accurate analysis with the well accepted half-amplitude threshold (half-amplitude) analysis (6) still requires 1) steady baseline; 2) signal-to-noise ratio $(i/\sigma) > 7$; 3) a single quantal current amplitude, with minimal overlapping currents; 4) sufficient temporal resolution to permit most openings or closings to reach a steady level of current; 5) clear separation of conductance-related fluctuations from gating transitions.

For many channel types, these conditions can usually be met, and the accurate analyses that have resulted have advanced our understanding of channel function over the past 15 years. Unfortunately, many circumstances exist in which such an analysis (particularly automated) fails to meet the criteria given above. It is particularly difficult to judge, ex post facto, the quality of the raw data used for an analysis when literature concerning such analyses present only the finished dwell time histograms and a handful of "representative" raw traces. An alternative format for summarizing single channel data is needed to demonstrate the suitability of data for half-amplitude analysis.

A broad range of channels are now known to violate the third condition cited above (quantal current amplitude) by displaying multiple levels of current when they are functioning. Often one open level predominates and is called the main conductance level of the channel. Smaller, current levels between the open and closed levels of a single channel are referred to as subconductance levels (sublevels). Na channels, for example, have been shown to have one or more sublevels in bilayers (16, 19) and in native membranes (10, 11, 12). These sublevels are occupied about 5% of the time that the Na^+ channel is open (12, 13).

Kinetic analysis is particularly difficult for channels such as these with multiple conductance levels. If the subconductance events are simply ignored (as for example by an automated analysis program designed to associate any crossing of the half-amplitude threshold with a gating transition), then significant bias may be introduced into the results. During a subconductance event the channel currents can remain close to the half-amplitude threshold for a prolonged period. Background noise is often sufficient during this time to cause multiple transient crossings of the half-amplitude threshold that lead to overestimation of the frequency of short openings and closings of the channel. Such problems can only be avoided if the subconductance events are accounted for by

Received for publication 25 December 1992 and in final form 8 February 1993.

Address reprint requests to Dr. Joseph Patlak.

0006-3495/93/07/29/14 $2.00

the analysis program, either by careful editing of the raw data to exclude any subconductance events, by fitting of each individual transition, or by idealization based on multiple half-amplitude thresholds between each possible level. None of these methods is entirely satisfactory, except perhaps for channels with very large currents. New methods are therefore necessary before the kinetics of channels with subconductance states can be addressed in a rigorous way.

Measuring the dwell times of the sublevel events themselves is particularly difficult in small conductance channels like the Na^+ channel. Even though i/σ is 8–10 for the main conductance state, the most common Na^+ channel sublevel has approximately 35% of this amplitude. The i/σ for distinguishing the sublevel from the baseline is therefore about 3, while that for distinguishing it from the main level is about 5–6. Such small signals provide no direct way to distinguish actual transitions into and out of the sublevel from random variations in the background noise. Additional filtering of the data improves the i/σ, but blurs the kinetics, making it difficult to distinguish brief sublevel events from the normal transitions between main and closed levels. Therefore, analysis of sublevel lifetime cannot be performed accurately using standard methods when the channel signals are small and when the sublevels are brief (both of which are the case for Na^+ channels).

I present a method here that can help to resolve the issues cited above. This method is an extension of a previous technique that was designed to measure accurately the current amplitudes of subconductance events. This "variance-mean" (12) analysis consisted of sliding a window, W time points wide, one point at a time over periodically sampled data. The mean and variance of the points in the window were determined at each sample point in the record. Whereas the previously published method constructed a "levels histogram" from those mean values that had variance estimates lower than the variance of the background noise, the method I present here simply assembles all the estimates into a three-dimensional histogram (mean versus variance versus number of events) for each window width. These histograms are thus an alternative representation of the *whole* data record, and they provide direct graphical display of the baseline stability and noise level, the amplitude of any conductance levels, the variance of each of these levels (open channel noise), the presence of poorly resolved flickers to other conductances, and the tendency for individual levels to interconnect kinetically. Gollasch et al. (7) recently demonstrated the utility of the such histograms for the analysis of single Ca^{2+} channels at physiological levels of charge carriers.

These mean-variance histograms (MV histograms) can also be quantitatively analyzed to determine channel kinetics. In the present paper, I also demonstrate that the volume of each low-variance region of such histograms is proportional to the integral of the dwell time distribution at that level for all times greater than the window width. By measuring the volume as a function of window width, an exponentially decreasing function is obtained. The time constants of this function are equal to those of the original dwell time distribution. Thus mean-variance histograms can be used to determine the time constants that describe the dwell time distributions for those levels with a dwell time constant longer than the sample interval, even when the i/σ is too low for other kinetic determinations. Other forms of quantitative analysis are also possible as well.

I present here the theoretical background for mean-variance analysis, the methodology used in generating and analyzing such representations of single channel data, the verification of its accuracy, even under conditions that make other forms of analysis unreliable, as well as practical lessons for working with such histograms.

MATERIALS AND METHODS

Generation of mean-variance histograms

Basic algorithm

The elementary operations used in generating a MV histogram are simple. For each sequential point in the raw data record, the mean current, $\langle I \rangle_t$, and sample variance, s_t^2, of that point and the preceding $(W - 1)$ points are determined, where W is the desired window width in sample points. Mean and variance are defined as follows.

$$\langle I \rangle_t = \frac{\sum_{i=1}^{W} I_{t+i}}{W} \tag{1}$$

$$s_t^2 = \frac{1}{W-1} * \sum_{i=1}^{W} (I_{t+i} - \langle I \rangle_t)^2 \tag{2}$$

This computation thus produces $(N - W + 1)$ pairs of mean-variance $(\langle I \rangle_t, s_t^2)$ values, where N is the total number of data points. For long data records, the storage and processing requirements of such real number pairs is prohibitive. Instead, they are binned and assembled into a two-dimensional histogram, of the form $MV(\langle I \rangle_t, s_t^2)$, where MV is the MV histogram.

The total number of bins used for each axis is arbitrary, as is the scaling. For computational convenience I have chosen to represent the MV histogram as an array of dimension 90×90 bins. Since variance estimates for most channel records can span up to four orders of magnitude, the variance axis uses log binning for maximum resolution. The following equations were used to calculate the bin number for each axis.

$$\text{Index}_M = \text{round}\left(\frac{\langle I \rangle_t}{\text{max}_M} * 72 + 72\right) \tag{3}$$

$$\text{Index}_V = \text{round}\left[\log\left(\frac{s_t^2}{\text{max}_V}\right) * \frac{90}{4} + 90\right] \tag{4}$$

Note the bias in this formula toward negative (inward) current, where 72 bins are available for negative mean currents, while 18 are for positive currents. Analysis of outward currents would require an appropriate adjustment of this bias.

When calculating in reverse from MV histogram bin number to the value at the midpoint of that bin, I used:

$$\langle I \rangle_{\text{mid}} = \text{max}_M * \frac{(\text{Index}_M - 72)}{72} \tag{5}$$

$$s_{\text{mid}}^2 = \text{max}_V * 10^{[(\text{Index}_V - 90) * 4/90]} \tag{6}$$

where $\langle I \rangle_{\text{mid}}$ and s_{mid}^2 are the current and variance values at the middle of the bin.

Finally, binwidth (used for fitting to theoretical curves) was calculated as follows.

$$\text{Binwidth}_M = \frac{\max_M}{72} \tag{7}$$

$$\text{Binwidth}_V = \max_V \left[*10^{\left(\frac{(\text{Index}_V + 0.5 - 90)*4}{90}\right)} - 10^{\left(\frac{(\text{Index}_V - 0.5 - 90)*4}{90}\right)} \right] \tag{8}$$

$$= \max_V *10^{(\text{Index}_V*4/90)} *1.0238*10^{-5} \tag{9}$$

Computational techniques

Calculation of new mean and variance estimates for each data sample, repeated for each desired window width, can be costly computationally. I used the following strategies to minimize the number of arithmetic operations needed to analyze a data segment: 1) Establishing a standard set of window widths for all computations. Eight MV histograms values were calculated simultaneously at window widths of 3, 5, 10, 15, 20, 30, 50, and 100 points. 2) Maintaining an internal, rolling array of the last 100 (equal to the widest window) data points. 3) Keeping the sum and summed square of the last W points for each window width W.

To calculate the new mean and variance, given a new data point, that point was first added into the previous sum, and its square was added into the summed squares for each of the eight window widths. The appropriate values for the oldest point in each window were subtracted out, and the rolling data buffer was updated with the new point (dropping the oldest). The mean and variance at each W are then easily determined from the new values of the sums and sum squares. This procedure, although less straightforward than simple computation of means and variances, eliminates the necessity of adding all W points (and their squares) for each new datum.

Baseline correction of raw data

The raw data used for constructing MV histograms must have a stable or a corrected baseline in order to attain maximum resolution of different current levels. I have used several techniques to correct for alterations in baseline, with the choice depending on the nature of the original data.

For data recorded at a steady potential, baseline changes were usually the result of slow drift. When such files (often megabytes long when sampled) needed baseline correction, I fitted a spline curve by eye to the baseline in a compressed display of the whole recording. The analysis program calculated the value of this spline at each time point and used this value to determine the shift necessary to zero the baseline.

When data resulted from voltage pulses, baseline correction was accomplished by using a combination of two techniques. First, blank traces were averaged and subtracted from the traces that had channel currents. Second, a levels histogram (12) was constructed from the data at window width 10. The baseline component of this histogram was fitted with a Gaussian curve in order to determine its midpoint. The data were subsequently shifted so as to zero the baseline.

Display of mean-variance histograms

MV histograms, like any other gridded data can be viewed in a number of different ways. The most practical for computer display purposes is to show the histogram as a two-dimensional area on the screen, with the amplitude of each bin encoded by the color displayed within a unit square. Although such color coding is acceptable using the standard 16-color displays of PC-type computers, 256 or more colors provide much better visualization of the data structures.

Since the number of entries in relevant bins can vary within a single MV histogram between one and millions, log scaling of the bin amplitudes is desirable. The formula that I used to encode color for the MV histograms displayed here was:

$$Color = \text{round}\left(Color_{\max} * \frac{\log(bin)}{\log(bin_{\max})} \right) \tag{10}$$

where $Color$ is the index within a gradation of colors (i.e. a spectrum), $Color_{\max}$ is the maximum value for the color index, bin is the number of entries within any bin, and $bin_{\max}$ is the largest bin in the displayed data.

Black and white renditions of the MV histograms were printed using the above formula where "color" encoded a shade of gray. Postscript-type printers gave the best performance for such prints. The color illustrations printed here were submitted to the publisher in electronic form.

Theoretical structure of the MV histogram

Distribution of mean and variance estimates in the MV histogram

As introduced above, mean-variance analysis measures the mean and variance values of a sliding window, W time points in width. Fig. 1 examines this relationship for idealized (no background noise, instantaneous transitions, only one channel active) single channel data. When that window is centered over a portion of the data in which the currents are steady (i.e., baseline, position A, or completely open, position D) the W data points are a sample from a population whose mean is the current level in that state and whose variance is 0. Such instances would plot to the lowest variance regions of a mean-variance histogram, as shown in the lower graph. If the window is positioned such that it contains a transition from one state to another, then the variance will be above zero, and the mean will be between the values for fully open or closed, as shown for positions B, C, and E. Thus a mean-variance histogram counts the instances in a long data record in which the window is aligned with data of constant, or changing current level.

In the case of real data, the points at any level will vary about their true mean due to instrumentation noise plus additional noise due to current flowing through the channel. The variance will never be as low as zero, in contrast to the idealized data shown in Fig. 1. Furthermore, since the W points within the window constitute but a small sample of the parent population, they will be dispersed about both the population mean and variance

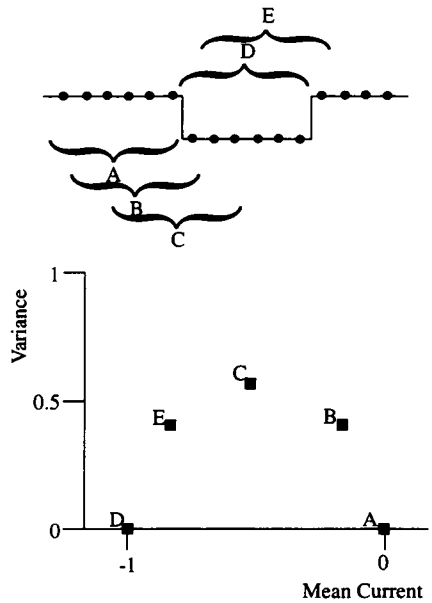

FIGURE 1 Relation between raw digitized data and plotted points on a mean-variance histogram. The top section represents the points sampled from a inward-going single channel event with single channel current of −1.0 units and no background noise. The curly brackets indicate five selected positions of a window of 6 points. The location of the mean-variance estimates at these positions is shown by the filled squares in the mean-variance histogram of the lower portion of the figure. The letters next to the symbols correspond to the indicated positions of the brackets.

values, even when the window is centered upon a steady level. For completely randomized data points, the distribution of the mean current estimates, $\langle I \rangle_t$, will be Gaussian, with variance given by:

$$\sigma^2_{\langle I \rangle} = \frac{\sigma^2_I}{W} \tag{11}$$

The estimates of sample variance, s^2_t, will have a χ^2 distribution with $W - 1$ degrees of freedom (Ref. 1, p. 83) so that

$$\frac{s^2_t}{\sigma^2_I} = \frac{X^2_n}{n} \qquad n = W - 1 \tag{12}$$

where σ^2_I is the population variance, and n is the number of degrees of freedom. In a MV histogram, this dispersion of the $(\langle I \rangle_t, s^2_t)$ estimates yields a two-dimensional "low-variance" region that characterizes each conductance level. By fitting to such low-variance regions in the data, one can accurately determine both the population mean and variance in that state. Furthermore, the *volume* of that low-variance region is the total number of instances in the data during which the entire window is fully within a steady current level. This volume is a function of both the window width, the frequency with which the level is entered, and the dwell time distribution for each occurrence of that level. I examine below the systematic relationship between the volume and these parameters.

Dwell time dependence of low-variance volumes for single channel currents

The number of counts, N, that each open, closed, or sublevel event contributes to its low-variance region depends on the length (dwell time) of the event, T_s. An event that is shorter than the window width must contain at least one transition (and thus has high variance), regardless of the window position. An event that has the same length as W will give a single instance in which all of the window is contained within the event (for example *position D* in Fig. 1); the variance will be low for this window position. Movement of the window by one point in either direction causes inclusion of a transition, increasing the variance, and thus moving the $(\langle I \rangle_t, s^2_t)$ estimates out of the low-variance region. Events longer than W lead to more than one entry, since the window is advanced one time point for each $(\langle I \rangle_t, s^2_t)$ pair. For events longer than W, subsequent $(\langle I \rangle_t, s^2_t)$ estimates will remain within the low-variance region until the final transition at the end of the event. The number of entries caused by a single event can thus be expressed as follows.

$$N = \begin{cases} 0 & \text{for} \quad T_s < W \\ T_s - (W - 1) & \text{for} \quad T_s \geq W \end{cases} \tag{13}$$

Single channel data consists of many such events with dwell times distributed according to some probability distribution. For example, the open times of a channel with a single open state, and with constant rates of transition that do not depend on the past history of the channel (Markovian kinetics) will be described by a monoexponential probability density function of the form:

$$\mathbf{P}(t) = \frac{1}{\tau} e^{(-t/\tau)} \tag{14}$$

where $\mathbf{P}(t)$ is the probability density for observing events of length t, and τ is the inverse of the sum of all rates leaving the state. The total number of low-variance counts, N_k, for a large number, k, open events at a window width W can thus be described by (see Appendix for derivation):

$$N_k = \kappa \int_W^\infty [t - (W - 1)]\mathbf{P}(t)\, dt \tag{15}$$

$$= \kappa\tau e^{\left(-\frac{(W-1)}{\tau}\right)} \tag{16}$$

where $\mathbf{P}(t)$ is as given in Eq. 2. N_k is therefore an exponentially decreasing function of W with time constant, τ, which is the same time constant as the original probability density distribution. When $\mathbf{P}(t)$ consists of the sum of exponentials, N_k will also consist of a sum of exponentials with identical time constants. Note, however, that the relative amplitudes of exponentials composing N_k is different from those in $\mathbf{P}(t)$, since N_k represents an integral function: Slow components will have a relatively larger amplitude in N_k.

The relationship between N_k and W in a given data set can be exploited to determine the time constants for individual levels in single channel data. The basic strategy is to assemble MV histograms from a set of data, then fit the low variance regions in order to determine their volume, N_k, at each window width. Finally, N_k is plotted as a function of W and fitted with one or more exponentials to determine the time constants at each of the levels. The integral nature of N_k also potentially affects the ability to fit curves to the $N_k(W)$ function. Since the points are not independent of one another, care needs to be taken in developing quantitative measures of goodness of fit, number of required exponentials, etc.

Note also that, when the exponential function best fitting the N_k vs. W plot is extrapolated to $(W - 1) = 0$, N_k will be equal to the number of events times the average duration of those events (see Appendix). In other words, the intercept for $W = 1$ is the total number of time points spent at the level. Dividing by the total length of the record, for example, gives the probability of being at that level during the recording.

Fitting the MV histogram

The sections above demonstrate that the shape, position, and volume of the low variance region in MV histograms have significance in the interpretation of single channel currents. The goal of this form of analysis is therefore to extract these meaningful parameters efficiently from the data. Although in principle one could fit the low-variance region with a unified, two-dimensional mathematical function, I found that such an approach was both slow and inaccurate in its fits. Instead, I have developed a more efficient strategy of fitting one-dimensional functions either to the mean currents at the level of the population variance, or to the variance distribution at the level of the population mean. These two approaches are described below using the example of the MV histogram (Fig. 2 *A*) derived from the data presented in Figs. 3 and 4 using a window width of 10 points.

For any single low-variance component of the MV histogram, the distribution of mean values should be Gaussian. A horizontal slice through the MV histogram at the level of the population variance should yield a simple histogram (directly analogous to an amplitude histogram or the previously published levels histograms) that represents this Gaussian. If multiple low-variance components (i.e. current levels) exist in the data, this simple histogram will consist of the sum of several different Gaussian components. The construction of these mean histograms from the MV histogram is illustrated in Fig. 2, *B* and *C*. These histograms were fitted with Gaussian curves of the form

$$N_m = A_m B_m P_m(\langle I \rangle) \tag{17}$$

$$P_M(\langle I \rangle) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(\frac{(\langle I \rangle - \bar{\mu})^2}{2\sigma^2}\right)} \tag{18}$$

where N_m is the expected number of counts in the bin, A_m is the total number of events in the component, B_m is the binwidth, $\langle I \rangle$ is the estimated mean current, μ is the population mean, and σ^2 is the variance of the distribution of $\langle I \rangle$ about μ (note that this is different from the population variance, though it is related to it under some circumstances). In the case of multiple components, the histograms were fitted with the sum of several such Gaussian components.

The fitting of mean-histograms was performed using a standard Levenberg-Marquardt algorithm (see Ref. 15, page 523). For each component the value for μ, σ^2, and A_m was optimized. The squared deviation between the actual number of observations in each bin and those predicted by a given set of parameters was weighted by the number of elements in that bin. The goodness of fit estimator was thus a χ^2 function, permitting the use of standard estimates as to the degree of fit improvement due to added components or the variability of the estimated parameters.

When multiple components were present, a separate mean histogram was taken at the level of the population variance *for each component*, as shown

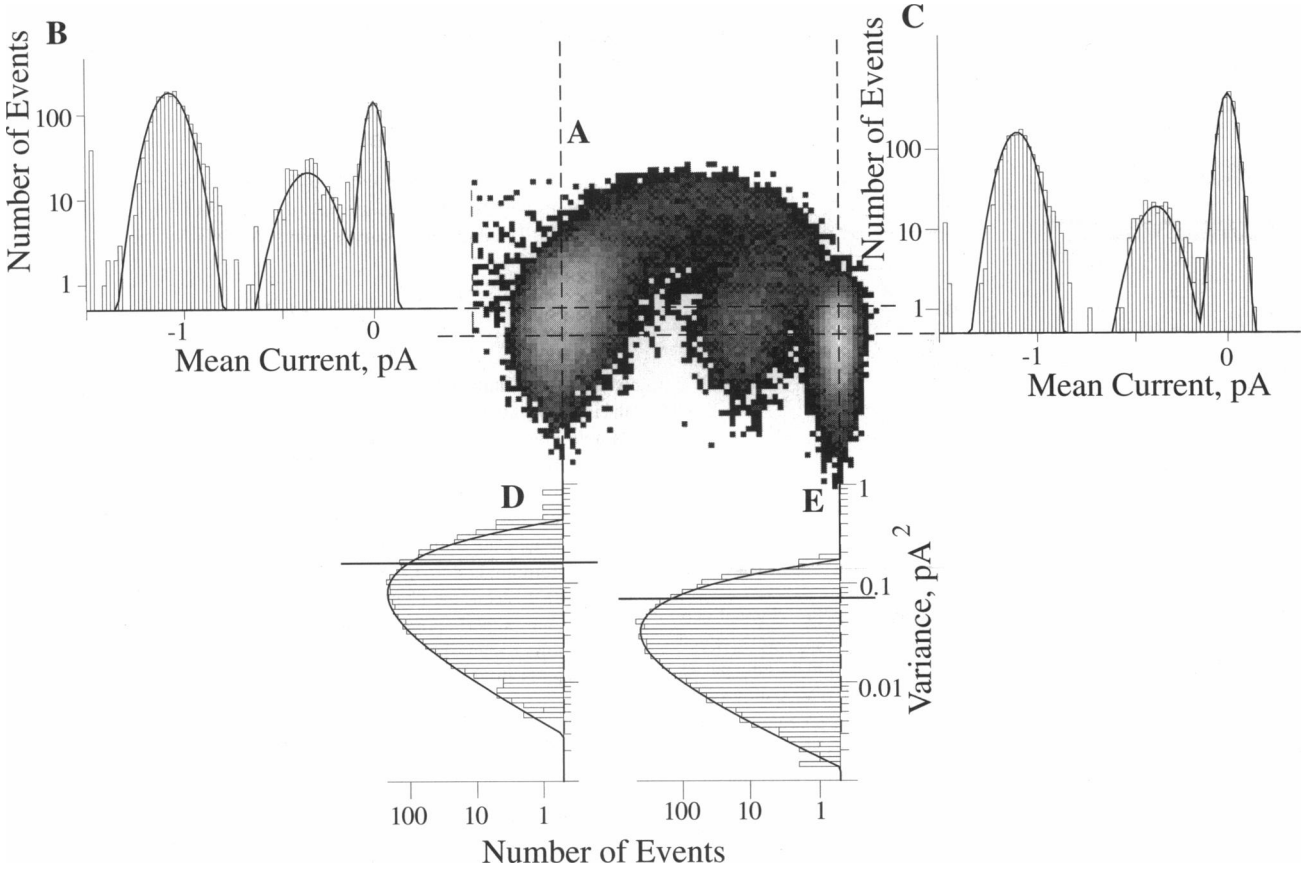

FIGURE 2 The method used to determine the best fitting volumes for each low-variance component of the MV histogram. *A* shows the raw MV histogram itself, with gray tones coding for the number of entries in each bin. The histogram was calculated from the data of Figs. 3 and 4 using a window width of 10. *B* and *C* show mean amplitude histograms made from the bins of the MV histogram that intersect the horizontal dashed lines extending from the base of each histogram. Note that since the open (left-most) component has a higher variance, the histogram is derived from a correspondingly higher level of variance. The histograms have each been fitted with the sum of three Gaussians, as shown by the solid line. *D* and *E* show variance histograms made from those bins of the MV histogram that intersect the vertical dashed lines. In each case the histograms were made at the level of the best fitting mean of that component. The solid curves indicate the χ^2 function best fitting the data. The horizontal solid line in each panel shows the maximum variance level that was used in each fit. The fits to the intermediate, subconductance level have been omitted for clarity.

in Fig. 2, *B* and *C*. This means, for example, that in a MV histogram with two low-variance components, two mean histograms would be constructed, and each would be fitted with the sum of two Gaussians. The fitting of the mean histograms was performed first using either preliminary (eye) estimates of the variance or the variance determined by the last cycle of fitting.

The variance distributions were next fitted by constructing a vertical slice through the MV histogram, as show in Fig. 2, *D* and *E*. One vertical slice was made for each low-variance component at the level of the population mean. The variance histograms usually consisted of a single component, possibly with some extra contamination from the high-variance points that form the arch in the MV histogram. As described above, this component is expected to have the form of a χ^2 distribution, and thus could be fitted with a function of the form:

$$N_v = A_v B_v P_v(s^2) \tag{19}$$

$$P_v(s^2) = \frac{n}{\sigma^2}[2^{n/2}\Gamma(n/2)]^{-1}\left(\frac{ns^2}{\sigma^2}\right)^{[(n/2)-1]} e\left(-\frac{ns^2}{2\sigma^2}\right) \tag{20}$$

$$n = W - 1$$

where N_v is the expected number of counts in the bin, A_v is the total number of events in the component, B_v is the binwidth, s^2 is the estimate of the variance, σ^2 is the population variance, W is the number of independent points leading to the estimate, n is the number of degrees of freedom, and G is the standard Gamma function (e.g. see Ref. 2, page 122). For each variance histogram, σ^2 and n were optimized using a Simplex algorithm (see Ref. 15, page 289) with goodness of fit and weighting determined as above for the mean histogram. As illustrated in Fig. 2 *E*, the fits were restricted to the bins between the lowest variance estimates and those representing twice the amplitude of the estimated population variance. This limitation served to reduce the influence of high-variance points on the parameters describing the low-variance region.

The value of n had to be fitted as an independent parameter for the variance histograms, because the data were filtered prior to sampling. When MV histograms were constructed from purely random numbers, the variance histogram was best fitted when n was the same as W. With filtered data, however, each point has some correlation with its neighbors, reducing the *effective* number of independent points leading to the mean and variance estimates. The magnitude of this diminution in the effective n depended upon the ratio of the sample to the filter frequency (oversampling). Adequate fits to the actual data could thus only be obtained when n was co-optimized.[1]

Since the mean histograms were constructed at the level of the population variance, and the variance histograms at the level of the population mean,

[1] Since the Γ function is defined only for integral and half-integral values of its arguments, optimization to real-numbered values of n were obtained by extending the Γ function. This extension was simply to interpolate linearly between the two nearest defined values of the function.

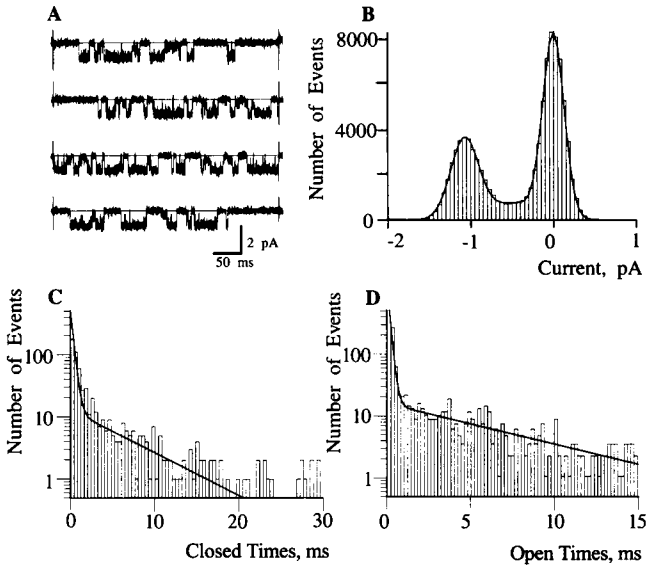

FIGURE 3 Standard analysis of single channel data. Na^+ channel currents were recorded from adult mouse skeletal muscle. The cells were treated with DPI 201-106, and the recordings were made at 0°C. 400 ms long pulses were given to −40 mV from a holding potential of −120 mV. *A* shows four representative traces out of 28 such bursts recorded at this potential. *B* is the standard amplitude histogram made from all the points during these 28 bursts, fitted with the sum of three Gaussians. The baseline had a mean of 0 ± 0. 128. The open component was fitted at −1. 07 ± 0. 17. The intermediate component had a mean at 0. 49 ± 0. 29. *C* and *D* show the closed and open time histograms from these data, along with their best fitting sum of exponentials. The closed time histogram consisted of 592 events, which were best fitted with time constants of 311 μs (66% relative amplitude) and 6. 3 ms. The open time histogram consisted of 561 events and was best fitted with 150 μs at 55% amplitude and 7. 2 ms.

the two fits were interdependent. The initial cycle of fits were made using eye estimates of these values, while subsequent fits used the values optimized from the previous cycle. Best results were obtained when two to three cycles of fit were performed on each low-variance component.

The information from the fits to the mean and variance histograms was then sufficient to determine the overall volume of the two-dimensional function. This volume was calculated as follows: The amplitude of any single bin, *mv*, in the two-dimensional histogram is given by

$$N_{mv} = V_{mv}B_vP_v(s^2)B_mP_m(\langle I\rangle) \tag{21}$$

where N_{mv} is the number of counts in the bin, V_{mv} is the total volume of the component in the three-dimensional structure. To determine this volume from the curves best fitting the mean and variance cross sections, note that these two orthogonal curves intersect at a single bin when $s^2 = \sigma^2$ and $\langle I\rangle = \mu$. For this bin,

$$N_m = N_v = N_{mv}. \tag{22}$$

Substituting for N_v and N_{mv},

$$A_vB_vP_v(\sigma^2) = V_{mv}B_vP_v(\sigma^2)B_mP_m(\mu) \tag{23}$$

$$V_{mv} = \frac{A_v}{B_mP_m(\mu)} \tag{24}$$

where σ^2 is the best fitting population variance, and μ was the best fitting estimate of the population mean for the component. These and other symbols are as defined above.

In some circumstances, it was not possible to obtain adequate fits to either the mean or the variance histograms. This occurred primarily when the number of entries into the component was very low or when the mean components had a non-Gaussian distribution. In such cases, the volume of the low-variance region was estimated by counting the total number of entries within the lower half of the low-variance region, as determined by fits to regions with similar variance properties.

Simulation of Single Channel Data

Channel kinetics were simulated by following a hypothetical channel through its possible states using random numbers to determine the times of transitions. The number of sample points, N_T, in any state was determined by Eq. 25 (3):

$$N_T = \frac{-\ln(U)}{kT_s} \tag{25}$$

where U is a uniformly distributed random number between 0 and 1, k is the rate constant for leaving that state, and T_s is the sample interval (seconds/sample point). When multiple pathways left a state, N_T was determined separately for each pathway, and the shortest value determined the transition taken.

Because the MV technique is sensitive to both the timing and trajectory of individual state transitions, it was not possible to mimic each transition with a prefiltered, stored, stereotypical transition as used in Magleby and Weiss (8). Instead, the channel kinetics were simulated at 10-fold higher sample rate, superimposed at this frequency upon normally distributed "white noise," then filtered digitally. The Gaussian-filtered points were then resampled at the appropriate rate.

Recording of Na^+ channel currents

The techniques used for the recording from mouse skeletal muscle were identical to those reported in Patlak (12). In brief, fibers were dissociated from the flexor digitorum brevis muscle of the mouse (BALB/c strain, Jackson Laboratory, Bar Harbor, ME). Muscles were dissected then enzyme-treated in physiological saline (PS) solution with collagenase and protease. The PS solution contained (in millimolar): 140 NaCl, 5 KCl, 0.2 $CaCl_2$, 0.2 $MgCl_2$, 10 HEPES (4-(2-hydroxymethyl)-1-piperazineethanesulfonic acid), 10 dextrose, titrated to pH 7.4 with NaOH. Subsequent to enzyme treatment the cells were soaked in "relaxing solution" which consisted of (in millimolars): 130 cesium aspartate, 5 KCl, 3 $MgCl_2$, 5 Na_2ATP, 10 HEPES, 30 dextrose, 0.1 EGTA ([ethylenebis(oxyethylenenitrilo)]tetaacetic acid), 4% Dextran T-500, titrated to pH 7.4 with CsOH. In experiments with lowered pH, a modified relaxing solution was used with an additional 10 mM of MES (4-morpholineethanesulfonic acid) buffer.

The patch recording setup was standard, consisting of a small chamber mounted on a temperature-controlled microscope stage, a Sylgard-coated patch pipette pulled from borosilicate glass, and a List EPC-7 (List Instruments, Darmstadt, Germany) patch clamp circuit. Patch currents were filtered at 2–3 kHz and sampled directly to a personal computer. During recording the cells were bathed in relaxing solution, which reduced their resting membrane potential to <5 mV. All voltages are thus given in absolute transmembrane differences.

Enantiomers of DPI 201-106 (4-[3-(4-diphenylmethylpiperazinyl)-2-hydroxypropoxy]-^{1}H-indole-2-carbonitrile), "DPI," were a gift of Dr. Scholtysik, Sandoz Inc., Basel, Switzerland. Stock solutions of the drug were made in ethanol, since it dissolved poorly in physiological saline. However, to avoid potential direct effects of the ethanol solvent on the membranes under study, approximately 2 μg of DPI's *S*-enantiomer were dried into the chamber before application of 1–2 ml of the cesium-relaxing solution, yielding a final concentration of much less than 2 μM (only a small fraction of the applied drug redissolved in the experimental medium).

RESULTS

Interpretation of MV histograms from Na^+ channel data

Standard analysis of Na^+ channel bursts

Fig. 3 illustrates the elements of a standard analysis of single channel data. In this case the recordings were of bursting Na^+

channels from mouse skeletal muscle. This experiment was part of a series designed to measure kinetics of Na^+ channels treated with DPI to remove inactivation. The temperature was held at 0°C. to maximize the temporal resolution of the recording. In addition, the pipette for this recording contained a low pH medium (pH 6.0), to obtain information about single channel conductance, the degree of open channel noise, the extent of sublevels, and the average open and closed times under conditions where H^+ ions transiently block the Na^+ channel. Fig. 3 *A* shows several representative traces taken during 200-ms pulses to −40 mV from a holding potential of −120 mV. The amplitude histogram constructed from 28 such pulses is show in Fig. 3 *B*. It shows a peak at the baseline with standard deviation of 0.13 pA, and another peak at 1.07 pA with a broader distribution. No other clear peaks can be seen, which could mistakenly be taken to indicate that sublevels are not significant.

Half-amplitude analysis should be valid on such data since the signal-to-noise ratio (measured as the ratio of single channel current amplitude to the standard deviation of the baseline noise) was about 8. The data were idealized by a standard half-amplitude algorithm, and representative traces were examined by eye for the accuracy of the idealization. Events shorter than two sample points were not counted as transitions in order to provide a consistent minimum threshold to use in correcting for missed events (6).

Histograms of the closed and open times that were produced by this idealization are shown in Fig. 3, *C* and *D*. The *solid curves* in the figures show the best fitting exponential functions using simplex optimization of the maximum likelihood estimator. The time constants of the open times histogram were 0.15 and 7.2 ms and were 0.31 and 6.3 ms for the closed times. Such kinetic and amplitude data would typically constitute a portion of a systematic study of channel function.

MV histograms from the same data

Fig. 4, *A–D*, show the MV histogram of this data set at four different window widths. Several important aspects of the data are available upon casual inspection: 1) The main open state has a higher variance than the baseline at all window widths. The open channel noise (0.02 pA^2) is thus immediately available. 2) A prominent region of subconductance activity is present, with amplitude of 0.36 pA, or 34% of the fully open current of 1.06 pA. 3) The low-variance region of the main open state is asymmetric, with a distinct tilt to the right as the variance increases. This is probably caused by a rapid, flickery component that is not well resolved at this sample frequency. 4) Clear interconnections between both the open and the closed states with the subconductance state can be seen, especially at higher window widths. Subconductance events can be entered from either the open or the closed states.

As stated under Materials and Methods, the low-variance regions (seen as the peaks in the three-dimensional representations in the *insets* of Fig. 4, *A–D*) are generated during intervals when the channel is at a steady conductance level. By determining the volume of these peaks, kinetic informaiton is derived about each of the states. Fig. 4, *E* and *F*, shows the best fit to the low variance regions of the MV histogram shown in Fig. 4 *B*. In Fig. 4*E*, the reconstructed two-dimentional regions are shown in isolation, while Fig. 4 *F* shows the difference between the actual data and the fitted volumes. With the exception of the fully open state, which shows some evidence of the asymmetry discussed above, the histograms are well fitted by the theoretical curves, as indicated by the random intermixing of negative (*blue-purple*) and positive (*yellow-red*) differences. Even the fully open state is well fitted by usual standards, since the largest differences are still less than 10% of the maximum bin heights (see, for example, the fit to the left-most component of the histogram in Fig. 2 *B*).

As discussed under Materials and Methods, the volumes of these low variance components are a function of the frequency of events, their dwell time distributions, and the window width. Fig. 5, *A*, *B*, and *C*, show the relation between this volume and window width for the closed, main open, and subconductance states, respectively. The best fitting exponential for the closed times has time constant of 16.7 ms and 0 intercept of 25,500. For the main state, these values were 7.8 ms and 32,000. Finally, the subconductance region had a time constant of 2.9 ms with an intercept of 5,300 points. This method indicates that: 1) Of the time spent in a conducting (i.e. nonclosed) state, 14% (5,300/32,000 + 5, 300) was at the subconductance level. 2) The dwell times at the subconductance level averaged only about a third of the dwell times in the fully open state. 3) Even though the time constant for the closed time was longer than either the open or sublevel states, the channel did not spend most of its time in that state. Rather, it appears to switch between open and sublevel (perhaps with a poorly resolved component of flickers leaving the open state).

Relationship between MV histogram results and those from half-amplitude

The results of the half-amplitude analysis differ from those produced by mean-variance analysis when applied to these data. In the face of such a difference, one must ask which of the two methods is giving the correct result. The answer appears to lie in the sensitivity of half-amplitude techniques to the ignored subconductance events in these data. The following paragraphs present the justification for this conclusion.

Fig. 6 shows a portion of the original data, selected to include several long subconductance levels. As can be seen in the figure, there are several instances where an opening would be registered erroneously during the sublevel intervals. The fast components in the fits in Fig. 3, *C* and *D*, could well have been due to such crossings. In addition, the extra events would have broken a single state into several shorter ones, decreasing the time constants.

Correcting for missed events would only degrade the situation further. If the fast components are simply ignored,

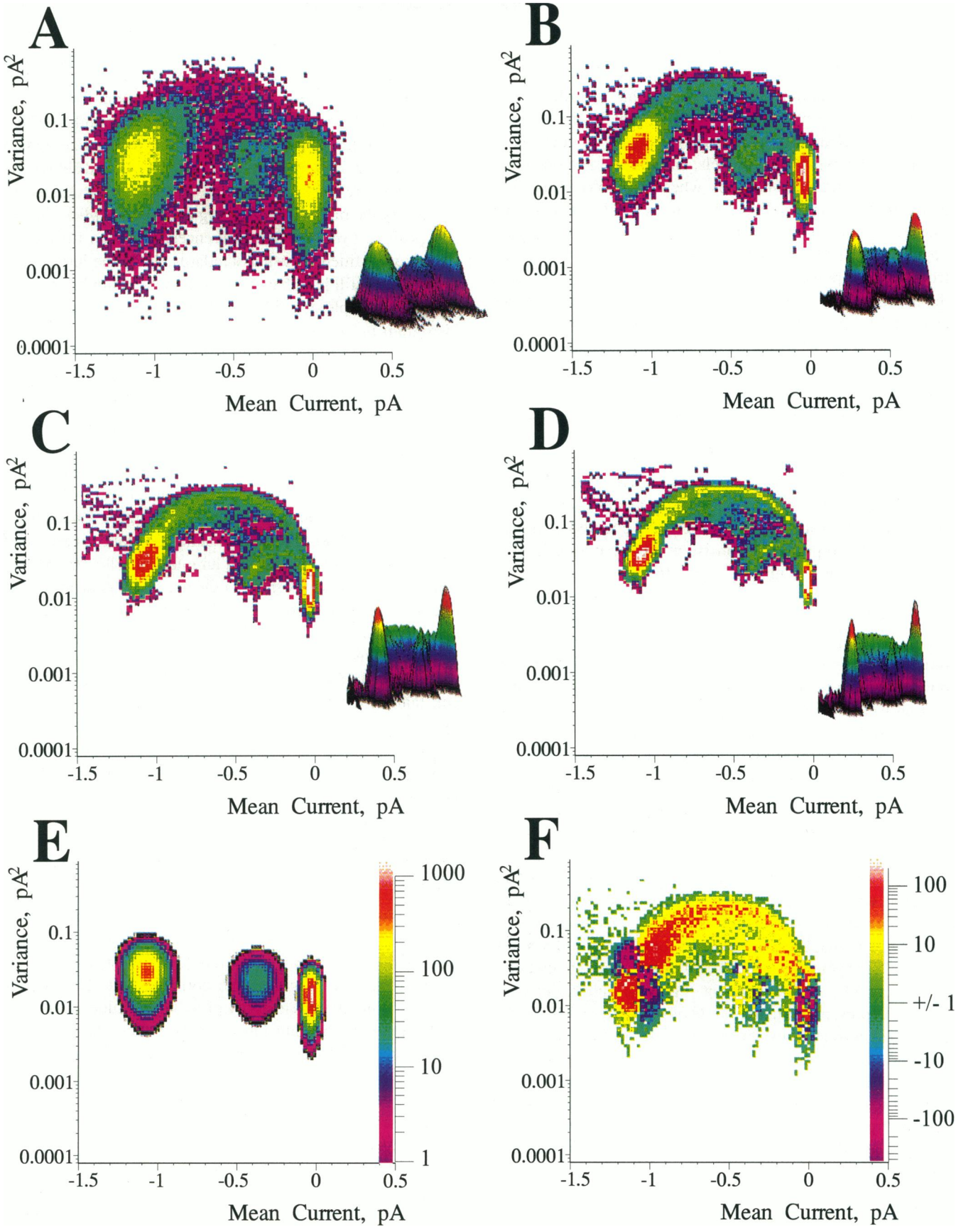

FIGURE 4 Mean-variance histograms constructed from the same data set as show in Fig. 3. Panels *A* through *D* show the raw MV histogram at four window widths, 5, 15, 30, and 50 points respectively. The color coding has been log scaled so that a bin with a single entry is black, while the bin with the maximum entry for the histogram is white. Intermediate entries are colored as for a standard spectrum. Panel *E* shows a representation of the best fitting

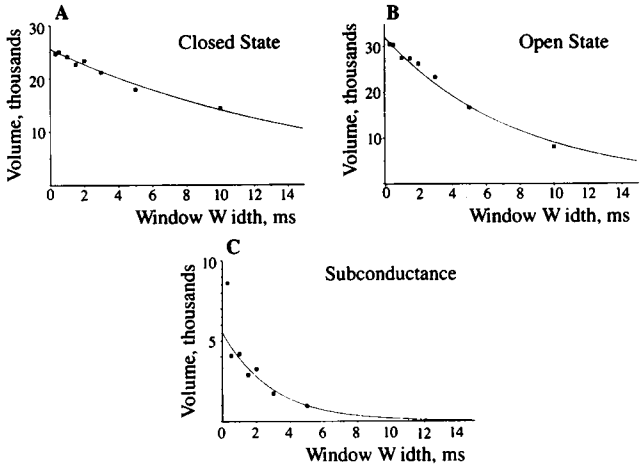

FIGURE 5 MV histogram low-variance component volume as a function of window width. Volumes were determined from the MV histograms shown in Fig. 4 (as well as other window widths not shown) using the fitting method discussed in the text. In each of the three panels, the measured volume has been plotted at the time corresponding to the $(W - 1)$ sample points. The best fitting single exponentials are shown as the solid lines. In *A*, the time constant for the closed component was 16. 7 ms. For the open component in *B*, the time constant was 7. 8 ms. *C* shows the fit for the major subconductance component, with time constant 2. 9 ms. The first point in *C* was excluded from the fit due to extensive overlap between the closed and sublevel components at this window width.

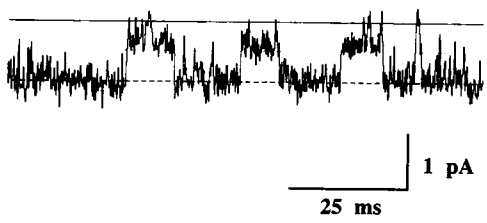

FIGURE 6 A high resolution view is shown of the currents during one of the pulses from the data set of Fig. 3. This segment was selected to illustrate the interaction between subconductance events and the half-amplitude threshold. The solid line is the zero current level, the dashed line is the fully open level of 1. 07 pA, and the dotted line is the half-amplitude threshold. Note that during the subconductance events the normal background noise is sufficient to cause many erroneous events to be registered.

then a slight decrease in the already foreshortened time constants would result. On the other hand, inclusion of these prominent fast components would lead to the conclusion that a large number of fast events were missed. Correction under these conditions would strongly decrease the values of the long time constants for both the open and closed states.

On the other hand, the MV histogram analysis has a demonstrable accuracy when applied to such complex data. The following section describes the testing of this technique using simulated data that was not correctly analyzed by half-amplitude methods.

Accuracy of MV histogram fits for determining kinetic parameters

The MV histogram analysis was tested for accuracy by fitting data produced from simulated single channel signals. The data were simulated with defined rate constants using a simple kinetic model. MV histograms were then constructed from the raw data and were fitted as described above. The technique provided accurate results over a wide range of kinetics and noise levels.

Fig. 7 shows results using simulated data that was generated with the simple kinetic model:

$$C \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} O \tag{26}$$

over a range of i/σ of between 20 and 2.5. k_1 and k_{-1} were both set to 1000 s^{-1}. Fig. 7 shows a plot of the rate constant values produced for the open and closed states as a function of i/σ, with the *dashed line* denoting the value used to generate the data. The method only starts to deviate significantly from the correct values below i/σ of about 3–5, and even at i/σ values of 2.5 the rate constants underestimated the correct value by about 66%. In contrast, rate constant estimates derived from half-amplitude algorithms degrade seriously below about 4, as shown by the *solid squares* in Fig. 7. At i/σ of 2 they are nearly 4-fold higher than the actual value.

Fig. 8 *A* shows an example of a more realistic simulation where the open/closed parameters changed as a function of membrane potential much as they do for the Na^+ channel. The kinetic model used for these simulation had the form:

$$\begin{array}{ccc} & k_1 & \\ C_s & \rightleftharpoons & O_s \\ & k_{-1} & \\ k_s \downarrow\uparrow & & \downarrow\uparrow k_{-s} \\ & k_1 & \\ C & \rightleftharpoons & O \\ & k_{-1} & \end{array} \tag{27}$$

where O_s was a subconductance state that was occupied approximately 7% of the time, and C_s was the same subconductance state with the voltage-dependent gates closed (i.e. it had 0 conductance and appeared as a closed state). The i/σ for these records was 7.5. Fig. 8 *A* shows the MV histogram constructed from these simulated data, while Fig. 8 *B* plots

functions to the three low variance regions in Panel *B*. *F* is also derived from the same panel, but color-codes the difference between the actual data in *B* and the fitted function in *E*. In this case, positive differences are shown by blue and purple tones, while negative differences are given by green, yellow, and red tones.

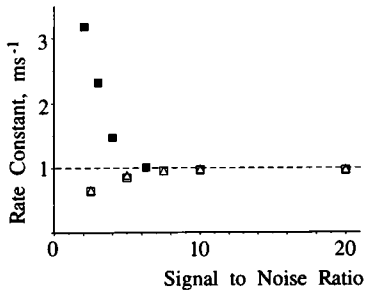

FIGURE 7 Rate constants derived from simulated data are shown as a function of the signal to noise ratio. Data were simulated at high temporal resolution using a simple open-closed model where both the forward and backward rates were 1000 s^{-1}, as indicated by the dashed line in the figure. White noise of varying amplitude was added to the channel signal, and the result was digitally filtered at 5 KHz and resampled at 20 kHz. Signal-to-noise was calculated as the ratio of the single channel amplitude to the standard deviation of the zero-current level (noise only). The estimates derived from Mean-Variance analysis are given by the open symbols, where the boxes are the opening rates and the triangles are the closing rates. These estimates were calculated from the equivalent of 50 seconds of data. For comparison, the solid symbols show rates derived from half-amplitude analysis taken from similar data sets. In these cases, the rates were determined by optimizing the fit of a single exponential to the open time histograms constructed from 500 events.

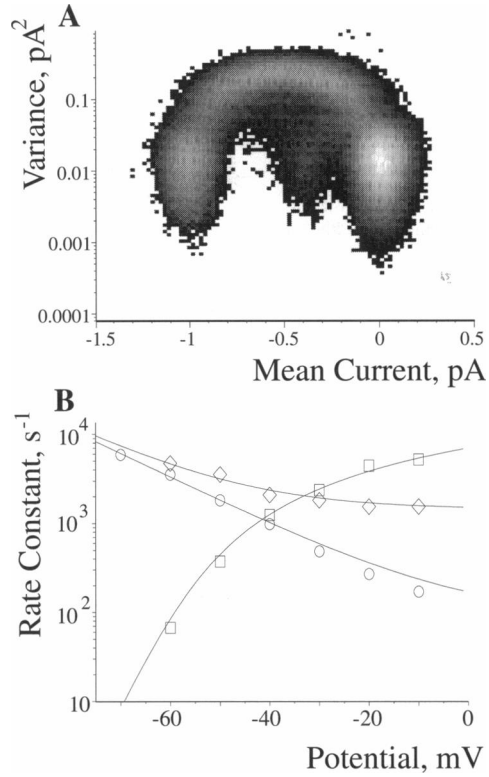

FIGURE 8 Rate constants derived from simulated data with a range of more complicated kinetics are shown. *A* shows the MV histogram (window = 10 points) derived from 50 seconds of data generated with methods identical to Fig. 7 using a signal-to-noise ratio of 7. 5. The effective potential for the recording shown was −50 mV. At each effective potential, the channel was simulated using the 4-state model shown in the text, where the rate constants were derived from equations previously used to describe the kinetics of bursting Na^+ channels. The ideal rate constants for leaving the open, sublevel, or closed states are shown as the solid lines in *B*. The open squares are the closed times measured from the MV histograms, the open circles are the open times, and the open diamonds are the sublevel times.

the predicted and measured values for the time constants, for each of the states. The curves used to calculate the original rate constants are show as *solid lines*, and the measured values are shown as *symbols*. As can be seen in the graph, the method yields accurate values over the full range of potentials. Therefore the MV histogram analysis appears to be a more informative and more accurate method for quantifying complex and noisy channel data.

A final simulation shows that the presence of subconductance states can lead to results like those shown in the half-amplitude analysis of actual data illustrated in Fig. 3. Fig. 9 *A* shows data simulated according to the kinetic scheme shown above, with rate constants adjusted to produce kinetics similar to that actually observed in the real data. Extra noise was added to the open and the subconductance states to make the simulated data as close to the actual channel currents as possible. Fig. 9 *D* shows the resulting MV histogram, which is directly comparable to those shown in Figs. 2 and 4, with the exception of the asymmetry in the open component. Analysis of the MV histogram produced kinetic rates very close to those used to generate the data: The open time constant was 7.8 ms, the closed time constant was 16.7 ms, and the subconductance time constant was 2.9 ms.

Fig. 9, *B* and *C*, show the closed and open time histograms produced by half-amplitude analysis of these data. A fast component was present in both histograms. The two time constants used to fit the open times histogram were 150 μs and 6.6 ms, which were very close to those measured for the real data. The time constants that best fit the closed times were 1.6 and 17.2 ms, both of which were slower than those fitting the real data in Fig. 3. In the case of the closed times, the very fast component registered in the analysis of actual data was missing in the simulated analysis. This is probably because the simulation did not include the fast, flickery closures that were not fully resolved by the MV histogram analysis, but which probably led to a component in the half-amplitude analysis. It is also likely that the intermediate (~2 ms) and the slow component merged into a single intermediate component (6 ms) due to the lower number of total events in the real data.

Despite the inadequacy of this simulation to reproduce all of the features of the original data, the tendency for introduction of extra components, and for the bias of the components that are indeed present is clearly illustrated. I therefore conclude from these analyses that for data with significant subconductance levels, the time constants derived from the MV histogram analysis are likely to be a better representation of the underlying kinetics than that produced by the standard half-amplitude analysis.

DISCUSSION

This paper describes a new method for the presentation and analysis of single channel data. The method is based upon my

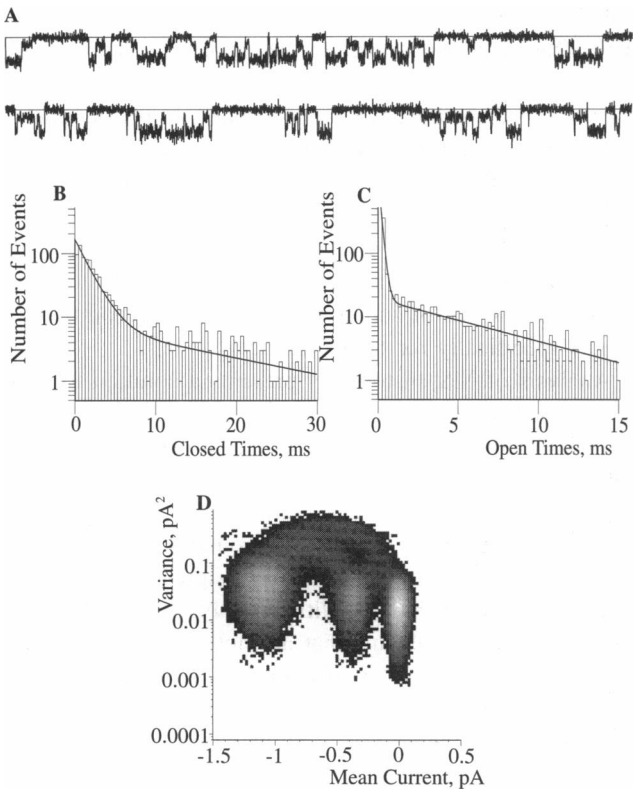

FIGURE 9 Data simulated to mimic the actual data shown in Fig. 3. *A* shows two segments of the simulated data, each 200 ms in effective length. The total amount of data analyzed was 30 seconds. In this case, the simulations were constructed with actual background noise (taken from blank traces in the actual recording), which was added to the simulated and filtered single channel traces. The 4-state model shown in the text was also used to simulate this data, with rate constants chosen to yield a mean open time of 7. 1 ms, a mean closed time of 16. 6 ms, and a mean sublevel time of 2. 9 ms. The open channel current was set to an average of 1. 07 pA, and the sublevel current to 0. 4 pA, but variability was allowed in these levels to give a spread in the low variance regions comparable to that of the actual data. *B* and *C* are the closed and open time histograms derived from these simulated data, along with their best fitting sum of exponentials. The closed time histogram consisted of 951 events, which were best fitted with time constants of 1. 6 ms (66% relative amplitude) and 17. 2 ms. The open time histogram consisted of 915 events and was best fitted with 150 μs at 64% amplitude and 6. 6 ms. *D* shows the MV histogram calculated from these same data for window width 10. Analysis of this set of MV histograms lead to estimates of 7. 8, 16. 7, and 2. 9 ms for the open, closed, and subconductance levels respectively.

earlier work (12, 13), but extends that work significantly in the type of information one can derive from single channel data that have multiple current levels or low signal to noise ratios. Although this presentation has been limited primarily to the analysis of a few sample recordings and of simulated data, these examples are representative of the complexities that many electrophysiologists face when trying to derive quantitative information from their data.

This method has proved useful for two aspects of data analysis. The first is as a qualitative graphic representation of the important characteristics of the channel currents during a recording. The second is as a quantitative tool with which to determine both the current amplitudes as well as the dwell time distributions of each component in the data. These two uses will be discussed separately below.

The MV histogram as a qualitative descriptor of single channel data

Single channel recordings produce large quantities of data. In general, one relies upon automated programs to derive specific kinetic or conductance information from the records. The two primary tools in this regard are amplitude histograms and half-amplitude idealization of the single channel currents. Amplitude histograms are generally used to determine the channel conductance, as well as to obtain an estimate of the open probability of the channels in the patch. In addition, they have been used to determine channel kinetics in cases where transitions are too fast to resolve (22). Standard amplitude histograms, however, are not very sensitive to either subconductance levels (which yield entries that mix with the normal transition points between open and closed), or to changes in the baseline during the recording. They produce only a coarse overview of the data.

Once such an overview has been produced, one generally proceeds to the idealization of the data using half-amplitude algorithms. Although such programs can produce high quality representations of the channel kinetics, they depend upon careful screening of all of the data input to them. This screening can be very tedious if done carefully by hand. Thus the temptation is great to simply turn the data over to the algorithm and check the resulting dwell time histograms for a "reasonable" appearance.

Unfortunately, half-amplitude analysis does not produce qualitatively inappropriate results when the data input to it is contaminated with extra levels, transitions that are not sharp, changing baseline, or low signal to noise. Rather, it gives a result much like one would expect from channel kinetics, with one or more roughly exponential functions characterizing the distributions. Therefore, it is difficult to determine from an amplitude histogram, dwell time distributions, and a few sample traces, whether the original data met all of the conditions necessary for quantitative analysis.

It would be highly desirable to have a more concise, yet more sensitive means by which single channel data could be summarized and presented. The mean-variance histogram is one such means. It is much like the standard amplitude histogram, yet its distribution of amplitude points is also sorted by variance, permitting clear separation of sublevel currents from sampled points in transition between the main levels. Furthermore, it changes its form in a characteristic way with channel kinetics, permitting the reader to gain an overview of this aspect of the data as well.

When preparing for a detailed kinetic analysis, an MV histogram can be a useful tool to verify the quality of the raw data. For example, it would be readily apparent from a single MV histogram that there was only one channel active throughout the whole recording, that the baseline was stable, that subconductance levels were not present, that the channel had little or no extra noise in the open state, and that the

overall signal-to-noise ratio was greater than about 7. Under these conditions the validity of a subsequent half-amplitude analysis could be firmly established.

Note also, that the MV histogram is an unbiased representation of the data. Unlike other techniques for current idealization (see below), the MV histogram presents all of the data without depending upon any assumptions as to their form. It is only the interpretation of the patterns produced within the MV histogram that depends upon an understanding of the underlying events. Thus the technique can be used to visualize data which has unknown or nonideal form.

Quantitative analysis of complicated single channel currents

A goal of most forms of single channel analysis is to derive kinetic information about the channel from the records of the channel's activity. For such analysis, one must determine the dwell times in each of the apparent states of the channel. In principle, this task should be relatively easy; just look for the time when the channel opens to more than half of its full amplitude, and measure the duration until it closes again to less than half amplitude. As discussed above, half-amplitude analysis yields accurate information about channel kinetics as long as a basic set of requirements about the data quality are met. Unfortunately, these requirements can often not be met in a number of interesting channel types, such as Na^+ channels, Ca^{2+} channels recorded in physiological divalent ion concentrations (7) or the absence of 1,4-dihydro-2,6-dimethyl-3-nitro-4-(2-trifluoromethylphenyl)pyridine-5-carboxylic acid methyl ester (Bay K8644), or mixtures of several types of K^+ channel in smooth muscle cells.

The two complications that have been of particular concern for my own data analysis in Na^+ channels have been signal-to-noise ratio and the presence of subconductance levels. Na^+ channel currents are relatively fast, and not particularly large. Of even more concern, however, are the presence of subconductance states in the records of Na^+ channels. In recordings of untreated Na^+ channels of mammalian skeletal muscle, the channel spent about 3–5% of its open time at one or more subconductance levels (12). In those channels treated with the drug DPI 201-106, this value climbed to about 5–6%, while in cases of acidification of the channel's environment, the channel can spend up to 20% of its open time in a subconductance level (14). Na^+ channels from brain and heart have equal or even greater presence of sublevels (10, 11, 18).

The mechanism for these subconductance events is unknown. In order to characterize them further it is necessary to have good measures of both their amplitude distributions, as well as the kinetics of the sublevel events. Furthermore, in order to distinguish between the kinetics of the channel's normal gating, and that caused by alterations in its conductance state (which may prove to be independent of the gating transitions), it is necessary to measure the dwell times in all the conductance states, not just the primary open and closed states.

Indeed, subconductance events cannot simply be ignored if accurate kinetic analysis is to result, even if one is only interested in the kinetics of the main states. Since they are present in many different types of channels, including voltage-dependent K^+, Ca^{2+}, and Na^+ channels, it is the responsibility of those analyzing such data to demonstrate either the absence of sublevels or to use a method that accounts specifically for their presence.

Analysis of MV histograms presents a solution to these problems. The technique, as presented here, permits not only an unambiguous identification of subconductance levels, but also the separate measurement of the dwell times and probability of occurrence for each of these levels. As an analysis method, it is relatively robust, and produces accurate results over a wide range of kinetics, even at signal-to-noise levels that would disrupt half-amplitude analysis. It is thus well suited for many problems that are not easily solved using other forms of analysis. Furthermore, the method deals efficiently with large volumes of data, since it does not require the operator to identify, fit, and/or check each individual transition between states.

The primary assumption of the quantitative analysis of MV histograms is that channels remain in discrete states, which appear as time periods with steady current and variance. These states can thus be recognized and quantified in the MV histogram as discrete regions with predictable form. Alternative approaches, which have been implemented by others, rely on different assumption about single channel signals. Events are taken to be separated by brief periods with rapidly changing current (20) or are assumed to be delineated by steps in the magnitude of the current (17). Chung et al. (4, 5) as well as Walsh and Sigworth (21) use a proposed model of the channel's kinetics to determine the most likely reconstruction of the original data. These, too, are all potentially valid approaches to the analysis of single channels, but they depend critically upon having single channel data that fits their picture of how channels behave. Complex channel data, with different degrees of open channel noise, gradual or variable rates of transition, or large background noise, may cause quantitative problems for these methods. In addition, like half-amplitude algorithms, they present only the finished analysis whose values are dependent on the degree to which the assumptions are met. In contrast, analysis of the MV histogram is relatively immune to complexities in the channel signals and starts from an unbiased representation of the data that presents qualitative information about the channel currents. The histograms are relatively easy to generate and to store. Once the raw data has been so processed, then subsequent methods for fitting can be applied as necessary to that data. Other methods that build their assumptions into the processing of the raw data require complete reanalysis if the assumptions change.

On the other hand, the MV histogram analysis technique is not without difficulties and limitations of its own. Its basic requirements are similar to other techniques: The data must have a stable or corrected baseline, and the channel currents must be large enough to resolve above the background noise. If the signal-to-noise ratio falls much below about 3–4, then the method tends to overestimate the time constants (for the simulated data presented here, the rate constants were about 66% the actual value for signal-to-noise ratio of 2.5). Instability in the baseline tends to have less effect on the measured time constants, but does decrease the ability to resolve small differences between levels or skew the best fitting Gaussians to inappropriate mean values.

Furthermore, the data must be filtered and sampled in a manner that is appropriate for both the method and the kinetic components to be measured. In particular, "oversampling" (sampling the original analog signal at more than twice the filter frequency) reduces the effective resolution at each window width, since each sample point is more highly correlated with its temporal neighbors. For half-amplitude analysis and display of single channel data, oversampling of at least 5–10 times the filter frequency is common. A better choice for MV histogram analysis is three to four times, although acceptable results can be obtained using higher values (the *patch* shown in Figs. 3 and 4 was sampled at 10 kHz and filtered at 2 kHz).

Once produced from sampled data, the MV histograms are only useful for fitting kinetic components within a finite range of values. Using a minimum window width of 3 points, it is impossible to resolve events shorter than 3 sample points, since they don't fit the basic definition of having reached a steady level of current and variance within this time. In order to have better temporal resolution using this method, it would be necessary to filter and sample the data at higher frequencies. Note, however, that fast events which do not settle into a low variance region are still represented in the MV histogram, and could possible be analyzed with another fitting scheme. There is also a practical limitation in the maximum time constants that can be resolved with the method as well. If most events are longer than the maximum window width (100 in the case of my analysis), then the decrement in number of low-variance entries as a function of window width will be small. Time constants measured from such data will have low precision. In such cases it would also be beneficial to resample the data at a lower filter setting and cut-off frequency.

The fitting method used here also presents some fundamental limitations. For example, I assume that the mean histograms consist of a sum of one or more Gaussians. This is an acceptable assumption if the components to be fitted are finite in number, discrete from one another, and are distributed due to random variation about a defined mean value. In some cases, however, a continuum of possible states may exist (12). In such a case, it may not be possible to resolve discrete components within this distribution. Rather, the MV histogram would display a low variance region which is a convolution of the mean current distribution of each state, and the probability distribution for observation of that state. Although it might be possible to separate these components, the present technique simply lumps them into a single broadly defined state. If the probability distribution is roughly Gaussian, this method is acceptable, but if not, then the fitting algorithm would have to be modified appropriately.

Even under optimal conditions, the method is only a tool to be used as part of the process of analyzing single channel data. It can neither provide direct insight into the kinetic scheme that actually produced the data, nor can it, by itself, provide direct measurement of all the rate constants of a complex scheme. These must be extracted using this and any other necessary tools as well as detailed knowledge of the abilities and limitations of those tools. With regard to the MV histogram analysis, it is not well suited to determining the temporal sequences within the data (e.g. the duration of bursts of openings), nor can it be used for conditional probability analysis in which the duration of an event is sorted based upon the properties of neighboring events (9). Finally, the user of the MV histogram method must be aware that it derives an integral of the dwell time distribution. As such, slow components will be overemphasized in their amplitude compared to fast components. In some cases, this may cause fast components to be overlooked despite their substantial presence.

Despite the limitations in the current implementation of the method, it will hopefully be a useful new tool for the analysis of single channel data. Indeed, many aspects of its utility remain to be explored. For example, would it be possible to fit the whole MV histogram, not just the low variance regions, in order to obtain an accurate account of all the kinetic activity? Could corrections be introduced for the overlap between low-variance and high-variance regions that degrade the fits at low signal-to-noise? Could one use the method to obtain the frequency dependence of open channel noise by systematically examining how the variance distribution changes with window width? These and other questions remain for future investigations to solve should the overall approach prove itself a useful tool.

APPENDIX

The formula for the distribution of volume versus window width is derived as follows.

For a defined channel state (i.e. open, closed, sublevel), the average dwell time in that state is the inverse sum of all the rates for leaving that state. Individual events will be distributed stochastically according to the probability distribution

$$\mathbf{P}(t) = \frac{1}{\tau} e^{-t/\tau} \tag{A1}$$

where τ is the inverse sum of the rates for leaving that state, t is the duration of any single event, and $\mathbf{P}(t)$ is the probability density for observing events of length t.

For any single event, the time spent during that event alone is t. For an ensemble of κ events with identical length, the total time spent during such events is κt. However, since the individual lengths are distributed according

to $\mathbf{P}(t)$, then the total time spent during κ randomly chosen events would be as follows.

$$\text{total time} = \kappa \int_0^\infty t\mathbf{P}(t)\,dt \tag{A2}$$

$$= \kappa \int_0^\infty \frac{t}{\tau} e^{(-t/\tau)}\,dt \tag{A3}$$

$$= \kappa\tau \tag{A4}$$

The result should be intuitively clear: The total time spent during κ events is the number of events times their average duration.

For the particular case of MV histogram analysis, we wish to know the total time spent during events of a particular state but following a dead time, w (defined as the window width − 1). For any one event, the time spent will thus be $t - w$, while for κ identical events it would be $\kappa(t - w)$. In the case of events distributed according to $\mathbf{P}(t)$, the total time spent during events longer than the dead time, N_k is:

$$N_k = \kappa \int_k^\infty (t - w)\mathbf{P}(t)\,dt \qquad w \equiv W - 1 \tag{A5}$$

To solve, let $y = t - w$ and substitute

$$N_k = \kappa \int_0^\infty \frac{y}{\tau} e^{-\left(\frac{(y+w)}{\tau}\right)}\,dy \tag{A6}$$

$$= \frac{\kappa e^{-w/\tau}}{\tau} \int_0^\infty y e^{-y/\tau}\,dy \tag{A7}$$

$$= \kappa\tau e^{-w/\tau} \tag{A8}$$

Thus N_k is a function of the window width, and it decreases exponentially with the same time constant, τ, as originally described the dwell time probability distribution. If $\mathbf{P}(t)$ consists of the sum of exponentials, so too will N_k. Note also, that when this function is extrapolated to $w = 0$ (i.e. window width of one), $N_k = \kappa t$, the total time spent at that state. When divided by the length of the recording, this gives the probability of being in that state at any time.

I thank Drs. Mark Nelson, Jurgen Hescheler, Maik Golash, and Thomas Kleppisch for helpful discussions and for their careful reading of the manuscript, and Mr. Darrell Hyde for assistance in preparing the color illustrations for publication.

The work was supported by National Institutes of Health grant AR-37606.

REFERENCES

1. Bendat, J. S., and A. G. Piersol. 1986. Random Data: Analysis and Measurement Procedures. John Wiley & Sons, New York. 566 pp.
2. Bevington, P. R. 1969. Data Reduction and Error Analysis for the Physical Sciences. McGraw-Hill, New York. 336 pp.
3. Blatz, A. L., and K. L. Magleby. 1986. Correcting single channel data for missed events. *Biophys. J.* 49:967–980.
4. Chung, S. H., J. B. Moore, L. G. Xia, L. S. Premkumar, and P. W. Gage. 1991. Characterization of single channel currents using digital signal processing techniques based on hidden Markov models. *Philos. Trans. R. Soc. London B Biol. Sci.* 329:265–285.
5. Chung, S. H., V. Krishnamurthy, and J. B. Moore. 1990. Adaptive processing techniques based on hidden Markov models for characterizing very small channel currents buried in noise and deterministic interferences. *Philos. Trans. Royal Soc. London B Biol. Sci.* 334:357–384.
6. Colquhoun, D., and F. J. Sigworth. 1983. Fitting and statistical analysis of single-channel records. *In* Single Channel Recording. B. Sakmann and E. Neher, editors. Plenum, New York. 191–263.
7. Gollasch, M., J. Hescheler, J. Quayle, J. Patlak, and M. Nelson. 1992. Single Ca channels from arterial smooth muscle with physiological concentrations of charge carriers. *Am. J. Physiol.* 263:C948–952.
8. Magleby, K. L., and D. S. Weiss. 1990a. Estimating kinetic parameters for single channels with simulation. A general method that resolves the missed event problem and accounts for noise. *Biophys. J.* 58:1411–1426.
9. Magleby, K. L., and D. S. Weiss. 1990b. Identifying kinetic gating mechanisms for ion channels by using two-dimensional distributions of simulated dwell times. *Proc. R. Soc. London B Biol. Sci.* 241:220–228.
10. Nagy, K. 1987. Subconductance states of single sodium channels modified by chloramine-T and sea anemone toxin in neuroblastoma cells. *Eur. Biophys. J.* 15:129–132.
11. Nilius, B., J. Vereecke, and E. Carmeliet. 1989. Different conductance states of the bursting Na^+ channel in guinea-pig ventricular myocytes. *Pflugers Arch.* 413:242–248.
12. Patlak, J. B. 1988. Sodium channel subconductance levels measured with a new Variance-Mean analysis. *J. Gen. Physiol.* 92:413–430.
13. Patlak, J. 1989. Measuring dwell time distributions of small subconductance levels using variance-mean analysis. *Biophys. J.* 55:A316 (Abstr.).
14. Patlak, J. B. 1990. The effect of pH on subconductance state frequency and duration in bursting Na^+ channels of skeletal muscle. *Biophys. J.* 57:A95 (Abstr.).
15. Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1986. Numerical Recipes: The Art of Scientific Computing. Cambridge University Press, Cambridge. 818 pp.
16. Recio-Pinto, E., W. B. Thornhill, D. S. Duch, S. R. Levinson, and B. W. Urban. 1990. Neuraminidase treatment modifies the function of electroplax sodium channels in planar lipid bilayers. *Neuron* 5:675–678.
17. Schindler, H., and V. Ph. Pastushenko. 1992. Statistical filtering of ion channel data. *Biophys. J.* 61:A248 (Abstr.)
18. Schreibmayer, W., H. A. Tritthart, and H. Schindler. 1989. The cardiac sodium channel shows a regular substate pattern indicating synchronized activity of several ion pathways instead of one. *Biochim. Biophys. Acta* 986:172–186.
19. Urban, B. W., E. Recio-Pinto, D. S. Duch, and M. Parnicas. 1987. Several conductance levels in steady state sodium channels. *Pflugers Arch.* 408:R311.
20. vanDongen, A. M. J. 1992. Transit: A new algorithm for analyzing single ion channel data containing multiple conductance levels. *Biophys. J.* 61:A256 (Abstr.).
21. Walsh., J. L., and F. J. Sigworth. 1992. "Hidden Markov" analysis ofionic-channel gating or, what to do with all those extra megaflops. *Biophys. J.* 61:A514 (Abstr.).
22. Yellen G. 1984. Ionic permeation and blockade in Ca^{2+}-activated K^+ channels of bovine chromaffin cells. *J. Gen. Physiol.* 84:157–186.