Topical Review

Computing the Field in Proteins and Channels

R.S. Eisenberg
Department of Molecular Biophysics and Physiology, Rush Medical College, 1750 West Harrison, Chicago, IL 60612

Received: 8 August 1995/Revised: 30 November 1995

Introduction

Predicting function from structure is a goal as old as molecular biology, indeed biology itself: the first glimpse of the anatomy of an animal, like the first glimpse of the anatomy of a protein, must have raised the question "How does it work?" The recent outpouring of protein structures (e.g., Creighton, 1993), often of enzymes in atomic detail (Singh & Thornton, 1995), has raised the same question again and again, and in a most frustrating way, at least for me, because the question is so often answered by description, not analysis. Of course, every biologist knows that a complex structure must be described before its biological function can be analyzed. But every physicist knows that description is not understanding. Indeed, if description is mistaken for understanding, vitalism will soon replace physics as the basis of biology.

Understanding enzymes is an important goal of modern biology because enzymes catalyze and control most of life’s chemistry. Enzymatic action depends on the diffusion of substrate and product, the conformation change of the enzyme, and the quantum chemistry of its active site. All must be analyzed in terms of physical models if enzymatic function is to be understood. The quantum chemistry of systems like enzymes is not well understood because enzymes are flexible structures of many atoms in a condensed phase, containing little empty space, and so are difficult to analyze with Schrödinger’s equation (Parr & Yang, 1989; Bader, 1990; Schatz & Ratner, 1993). The movements of the substrate, product, and enzyme must also be understood, presumably as a result of electromechanical models, like Langevin equations (Kramers, 1940; Gardiner, 1985; Hynes, 1985, 1986; Hänggi, Talkner & Borkovec, 1990; Fleming & Hänggi, 1993). Movements of substrate and product have been studied in this tradition, but movement of the protein (i.e., conformation changes) have rarely been connected to the physics that govern them.

Conformation changes are widely used to explain biological function (Albers et al., 1994) and are undoubtedly common and important mechanisms, although perhaps not quite as common as sometimes assumed. I suspect that the idea of conformation change was originally introduced to seek "... the ultimate source of the autonomy, or more precisely, the self-determination that characterizes living beings in their behavior" (Monod, 1972, p. 78). The identification and description of these conformation changes is one of the main tasks of molecular biologists. But the description of the conformation change (as allosteric, or whatever) is not the same as a physical understanding. Without a physical model of catalysis and conformation change, understanding of enzymatic function is not possible. The physical model of conformation change will use the language of electrodiffusion, because the atoms that move inside the protein are usually charged, and always diffuse, according to Langevin equations much like those that describe ionic motion in solutions.

A physical model of electrodiffusion is feasible and various forms have been used to analyze transport processes in biology for more than a century. Recently, the

---

Correspondence to: R.S. Eisenberg

1 Linnaeus must precede Darwin; or rather "Voyage of the Beagle" must precede "Origin of Species.

2 in a chapter entitled 'Microscopic cybernetics.'
nonlinear Poisson-Boltzmann equation (\(PB_n\) for short), has been widely used to analyze (the average properties of) diffusion in enzymes (Davis & McCammon, 1990; Honig & Nicholls, 1995; Hecht et al., 1995, is particularly important because it provides direct experimental verification of the theory; also: Warwick & Watson, 1982; Klapper et al., 1986; Gilson, Sharp & Honig, 1988; Davis et al., 1991). \(PB_n\) is a major contribution to our understanding, but \(PB_n\) is not a theory of enzymatic function, because it describes only electrodiffusion—and that at equilibrium. It is not a theory of catalysis or conformation change. The question then arises whether a significant biological function can be understood just from an understanding of electrodiffusion?

Electrodiffusion is a main determinant of the transport of substances across biological membranes. Understanding membrane transport has been an important goal of biology for more than a century because it governs so much of life. In particular, membranes—in their role as gate-keepers to cells—are responsible for signaling in the nervous system; for coordination of the contraction of skeletal muscle and the heart, forcing its muscle to function as a pump. Membranes contain receptors or effectors for many drugs and natural substances that control the life of cells.

The classical analysis of biological transport as just a form of diffusion has, however, been only partially successful (see Hille, 1989; Läuger, 1991, for a modern perspective). Such analysis usually assumed that diffusion occurred in systems of fixed structure, in particular of fixed cross sectional area, but we now know that the area for diffusion is modulated and controlled in most biological systems by the opening and closing of pores, namely channels in membrane proteins (Hille, 1992). Indeed, determining the number of open channels and their modulation is a main task of physiologists nowadays (Alberts et al., 1994; Hille, 1992).

The Open Channel

Although modulation of opening is a complex process, and opening itself may be as well, a channel, once open, forms a simple well-defined structure of substantial biological importance. An open channel should be “wholly interpretable in terms of specific chemical [or physical] interactions . . .” (Monod, 1972, p. 78). The age old question “How does it work?” should be easier to answer when ‘it’ is an open channel, than when it is anything else, at least any other class of proteins of such general significance.

The starting place\(^3\) for a theory of open channels is a theory of electrodiffusion rather like that used previously to describe membranes. The theory uses Poisson’s equation to describe how charge on ions and the channel protein create electrical potential; it uses the Nernst-Planck equations to describe migration and diffusion of ions in gradients of concentration and electrical potential. Combined,\(^4\) these are also the “drift-diffusion equations” of solid state physics, which are widely, if not universally used to describe the flow of current and the behavior of semiconductors (Ashcroft & Mermin, 1976) and solid-state devices, like transistors (Sze, 1981; Selberherr, 1984; Rouston, 1990; Lundstrom, 1992). The drift-diffusion equations describe the shielding or screening of permanent or fixed charge whereby the ions in the ionic atmosphere in and around a (channel) protein help determine the potential profile of its pore, a phenomena long known to be biologically important (Frankenhaeuser & Hodgkin, 1957; McLaughlin, 1989; Green & Andersen, 1991). Mathematical difficulties have been limiting, however, and so attention has usually been focused on the ionic atmosphere at the surface of the membrane or ends of the channel, and not the co- and counter ions within the channel’s pore.

Many theories focus on systems at equilibrium in which all fluxes are essentially zero. In the latter case, the PNP equations reduce to the (one-dimensional) \(PB_n\) equations. A great deal of important work has been done on \(PB_n\) and in some ways \(PB_n\) is the most physical theory of proteins now available; nonetheless, it is of limited use in understanding the open channel because the natural biological function of channels occurs only away from equilibrium (as does the biological function of most enzymes!). The biological function of both channels and enzymes is usually flux.

Significant fluxes flow even at the reversal potential of a typical, imperfectly selective channel. Fluxes are zero only in a perfectly selective channel at its reversal potential, which is indeed only then an equilibrium potential. A generalization of the \(PB_n\) equations is needed to predict flux and PNP is one such generalization. Interestingly, the potential profiles of semiconductors have

---

\(^3\)Electrodiffusion in biology occurs in a phase with little empty space, in which friction and interactions predominate and so the starting point for a theory of channels should be a theory of liquids or condensed phases. As Berry, Rice and Ross put it in their textbook (1980, p. 844, emphasis added): “. . . the principle difference between a dilute gas and a liquid . . . [is that] in a dilute gas a typical molecule is usually outside the force fields of all other molecules and only occasionally in the force field of one other molecule (binary collision); whereas in a liquid a typical molecule is usually within the force field of, say, 10 nearest-neighbor molecules and is never completely free of the influence of other molecules.” The starting place must not be a theory of gas phase kinetics, like Eyring rate theory, because it is likely to miss the important effects of interactions (i.e., friction), that dominate the properties of liquids (Purell, 1977; Berg, 1983) but not gases.

\(^4\)We call the combination PNP to emphasize the importance of the electric potential and the analogy with solid state physics.
a quite restricted repertoire of behavior at equilibrium. Semiconductors cannot be transistors, switches, or amplifiers in the absence of current, or even in the presence of just small currents, and so an equilibrium theory can be expected to give only limited insight into the repertoire of natural functions of channels, even if it could calculate one.

Many theories of ion movement include only some types of charge (e.g., permanent charge is usually ignored in the electrochemical literature cited below). But all charge is likely to have global effects, at least judging from work on charge at the ends of channels (Green & Andersen, 1991). Thus, the whole system, containing all types of charge and flux, has to be analyzed if the biological function of the open channel is to be predicted successfully from its structure.

**PNP Theory**

The mathematical difficulties in the analysis of the full system have been largely overcome because a rapid and accurate (numerical) integration procedure is now available (even though a qualitative analytical theory is not known). PNP theory can now predict the current through an open channel given its structure and distribution of fixed (i.e., permanent) charge in a few seconds of computer time. Indeed, once the structure of the open channel is known, and thus the distribution of its permanent charge, along with diffusion constants, PNP theory predicts its properties—the fluxes and currents through it—in all experimental conditions of varying concentrations and transmembrane potentials. Figure 1 illustrates the several components of charge in PNP and Fig. 2 shows how the structural data (i.e., permanent charge, length and diameter of the channel; and the dielectric constants) plus experimental conditions (i.e., potentials and concentrations in the baths) determine the output of the theory, namely current-voltage relations. PNP theory is the numerical procedure described in the Appendix, which solves Poisson’s differential Eq. (A1) simultaneously with the integrated Nernst-Planck equations (A7) using the boundary conditions (A2)–(A5).

Preliminary work shows that PNP theory—using a few (say, five) adjustable parameters—fits a wide range of data, taken from many solutions, that is difficult to fit with traditional models (Franciolini & Nomber, 1994a,b; Kienker, DeGrado & Lear, 1994; Chen et al., 1995a,b; Kienker & Lear, 1995). PNP automatically predicts a wide repertoire of behavior, because it is nonlinear and the potential profile in the channel and its pore (which is an output, not assumption, of the theory) changes significantly with experimental conditions. A post-hoc qualitative understanding of this behavior is possible in many cases, but PNP is a mathematical theory, a set of coupled nonlinear differential equations, describing interactions arising from all types of charge, and flux. It is rarely possible to predict the behavior of such systems before they are solved, and it is not always possible to rationalize the behavior of such systems post hoc, in a few words, particularly if several terms in the (mathematical) solution, or types of charge (some positive, some negative) are significant.

PNP theory fits a wide range of data because shielding usually has global effects, spreading across the entire space.
channel. If solutions or membrane potential are changed, the ionic atmosphere of counter and co-ions in and near a channel (and thus the potential and contents of the channel’s pore) change, as they must, if potential and concentrations simultaneously satisfy Poisson and Nernst-Planck equations. Our calculations show changes of potential of several \( k_B T / e \) in most locations when typical solution changes are made. In loose terms, we can say that small changes in net charge in and near the channel make significant changes in potential and even bigger changes in flux; potential is a sensitive function of net charge, and flux depends exponentially on the potential profile \( \Phi(\zeta) \).\(^6\) The ionic atmosphere and shielding are major determinants of a channel’s properties.

PNP theory is significant as a theory of an important biological phenomena—open channel permeation—arising from a simple physical mechanism, electrodiffusion. But it is also important in the more general context of proteins and enzymes as well, because PNP theory shows by implication, if not derivation or proof, that any property of a protein will be strongly influenced by changes in the electric field, whether gating of a channel, mediated or active transport of a permease, conformation changes in an enzyme, or catalysis itself. Indeed, we suspect that many of these processes will be dominated by the electric field and its change in shape. PNP adds another example\(^7\) to those already known, in which the electric field dominates the biological function of a protein (see also Warshel, 1981; Warshel & Russell, 1984; Davis & McCammon, 1990; Warshel & Aqvist, 1991; Honig & Nichols, 1995).

### Electrodiffusion in Semiconductors, Solutions, and Channels

The fundamental physical process in transistors and semiconductors is the migration and diffusion of charged quasi-particles—holes and electrons—in electric fields and gradients of concentration, just as the fundamental process in channels is the diffusion of ions, and perhaps quasi-particles, as well. In electrochemistry and semiconductor physics, the (mean) electric field is usually described by Poisson’s equation\(^8\) (that specifies how charge creates potential)

\[
-\varepsilon H, eF \frac{d^2 \varphi}{dx^2} = \varepsilon P(x) + \frac{e}{\varepsilon} \sum C(x) \]

\[
\frac{d}{\varepsilon} \left( \nabla \cdot \left[ \begin{array}{c} \nabla (1 - \varepsilon d) - \varphi(x) \end{array} \right] \right)
\]

\[
\text{Deviation From Constant Field}
\]

Symbols and terms are discussed in the Appendix.

Electrodiffusion is usually described by the Langevin equation (that describes individual trajectories: Kramers, 1940; Gardiner, 1985; Eisenberg, Klosek & Schuss, 1995) or the Nernst-Planck equations (that describe the probability density function of these trajectories). Here, the Nernst-Planck equations are written in the integrated form we have found most useful.

\[
J_f = D_f C(L) \exp(\zeta \phi(\zeta) d_x) \int_0^d \exp(\zeta \phi(\zeta) d_x) \right)
\]

\[
I = \pi \zeta^2 \sum_j \zeta_j |J_f|
\]

These equations can be written exactly as rate equations for any potential profile \( \Phi(\zeta) \) (see Eqs. (A9) and (A10) in the Appendix), with rate constants given by specific statistics of the underlying random atomic trajectories (Eisenberg, Klosek, & Schuss, 1995). In this way, traditional rate constant models can be connected to PNP theory; the important difference being that the rate constants of PNP theory are coupled to the entire (nonlinear) system. They depend on the entire potential profile \( \Phi(\zeta) \), which in turn depends on all experimental conditions (for example, the imposed concentrations in the baths \( C(L) \) and \( C(R) \) and the imposed transmembrane potential \( V_{\text{app}} \)).

The Nernst-Planck equations are fundamentally nonlinear because the conductance of ionic solutions depends on concentration.\(^9\) The Nernst-Planck equations are fundamentally coupled because they depend on the potential profile \( \Phi(\zeta) \) which in turn depends on everything else in the system, through the Poisson equation. Thus, as the concentration changes, the migration of ions

\(^6\) Eq. (2) illustrates the dependence of flux on potential and concentration within the channel. Eq. (A7) illustrates the dependence of concentration on potential. Understanding the dependence of potential requires numerical solution of the full set of PNP equations, see Appendix, p. 30 and footnote 15.

\(^7\) A particularly convincing and biologically relevant example, in the parochial opinion of channologists like me.

\(^8\) Variables, dimensions of variables, brief explanation and references can be found in the Appendix.

\(^9\) This essential point is easily obscured by the several forms of equations and the plethora of parameters used to describe migration—believe it or not, conductance, conductivity, mobility (conventional and absolute), and diffusion constant are all used in elementary texts (Brockis, & Reddy, 1970; Bard & Faulkner, 1980) and advanced treatments (e.g., Justice, 1983).! The essential point is that the conductance of a given volume of solution depends on the concentration of ions in the solution and the conductance appears as a coefficient of a derivative in the Nernst-Planck differential equations. Concentration varies over a wide range in most systems and so linearized approximations cannot describe their qualitative properties.
changes, even if everything else is constant. The Poisson equation is nonlinear (in the present context) and coupled because the net charge of ions $e\Sigma z_i C_i(x)$ is one of the source terms in the equation, but that term depends on the concentration of ions $C_i(x)$, which is an output of the Nernst-Planck equations (see Eq. (A7) of the Appendix), which in turn depend on the potential profile $\Phi(x)$ as we have just stated.

The nonlinearity of the coupled Poisson and Nernst-Planck equations allows a richness of behavior that we use every day, given the role of solid-state electronics in our technology and economy. Transistors are semiconductors designed to do specific jobs, to have nonlinear properties that arise in part from the nonlinearity of the Nernst-Planck equations, in part from the nonlinearity of their coupling to the Poisson equation discussed later in this paper. The distribution of permanent charge is chosen by the designer of the transistor to create the shape of the electric field he wishes, thereby making an otherwise uninteresting homogeneous lump of pure silicon into a switch, an amplifier, a detector. An entire computer can be built solely out of a network of transistors that obey the Nernst-Planck (and Poisson) equations. All written knowledge, and all mathematical operations, can be stored or executed by a computer, and so mathematical solutions to the Nernst-Planck (and Poisson) equations can have a rich range of behavior.

The PNP equations can themselves produce the microscopic cybernetics thought by some philosophers of science to be characteristic (or defining) of life (Monod, 1972, p. 68–80), even in the absence of the allosteric conformation changes they had postulated. Allosteric conformation changes arise, of course, from the movement of charged particles, described by equations much like PNP, and so both electrodiffusion and conformation change can themselves have the rich nonlinear behavior of transistors, which indeed is richer than the nonlinear behavior of traditional (rate theory) models of allostery (Monod, 1972). Of course, no one yet knows if channels actually perform these cybernetic functions by changing the shape of their electric field, and if such functions are important for the life of the animal as a whole, as likely as it seems, given their importance in semiconductor technology.

Traditional Models of Open Channels:
Rate Constants

Traditional models usually describe the structure of channel proteins as a distribution of potential, 'a potential of mean force' which in turn determines rate constants that describe open channel permeation, or gating (Andersen & Koepp, 1992; Hille, 1992) or changes in conformation in general (Hill, 1977; Walsh, 1979; Hill, 1985). Such models are both too flexible and too rigid, in my opinion. They are too flexible because the operational definition of 'state' (in a condensed phase dominated by friction, like a protein and its pore) is often vague and indeterminate: states and adjustable parameters can be added too easily to make a model fit too wide a range of data. They are too rigid because the rate constants of barrier models are nearly always assumed to be independent of concentration of ions (or substrate) and that is nearly impossible, given the importance and variability of shielding in any system with mobile charge.

A theory can easily be checked to see if it is consistent with modern ideas of condensed phases (see footnote 3). Flux in a condensed phase always depends on friction; it is wise then to check that the theory being used displays the dependence of flux on the diffusion constant or a derived parameter. Many do not (Hill, 1977; Walsh, 1979; Hill, 1985; cf. theories of condensed phases, Chandler, 1978; Hynes, 1985, 1986; Berne, Borkovec & Straub, 1988; Fleming & Hänggi, 1993). Unfortunately, nearly all rate constant theories of gating and open channel permeation do not display the role of friction (the large literature in these fields can be reached through the references in Hille, 1992; Andersen and Koepp, 1992; Läuger, 1991, and Eisenman and Horn, 1983).

A traditional theory can also be checked to see if it is consistent with the laws of electricity: it is consistent if and only if the potential profile satisfies Poisson's equation (and the rate constants in it are computed from the potential profile, cf. Eqs. (A9) and (A10) of the Appendix). In one-dimension this means the second spatial derivative (the curvature) of the potential must equal\textsuperscript{11} (at every location) the sum (at that location) of all charges in the model, including partial charges found on nearly every atom of a protein (see pp. 8 and 17). If the charges in the model do not add up to the second derivative, the model is inconsistent and incorrect. This check has to be repeated in each experimental situation (e.g., for each set of bath concentrations and potentials) and at each location in the system. Rate theories of gating and permeation (loc. cit.) never satisfy this condition (to the

\textsuperscript{10} An excess of flexibility [in a theory] may well turn power into weakness. For a theory that explains too much ultimately explains very little. Its indiscriminate use invalidates its usefulness and it becomes empty discourse. Enthusiasts and popularizers, in particular, do not always recognize the subtle boundary that separates a heuristic theory from a sterile belief; a belief which, instead of defining the actual world, can describe all possible worlds." (Jacob, p. 22, 1982; cf. Monod, 1972, p. 69 et seq.). If Jacob and Monod had realized the relation of their work to that of Bardeen, Britain, and Shockley; if they had known that rate constants can only be derived from nonlinear theories like PB, or PNP that have a wide range of 'oriented, coherent, and constructive' behavior, perhaps the literature of allosteric models would be smaller. It presumably would not include so many papers that use underdetermined rate and state models to describe complex experimental systems of unknown structure.

\textsuperscript{11} Ignoring units, for the sake of simplicity in writing.
best of my knowledge), when used to describe a range of experimental conditions, since they never compute the rate constants from solutions of PNP or other self-consistent field equations. They usually assume rate constants independent of bath concentration and transmembrane potential. Many diffusion theories of permeation do not satisfy this check either because they describe a range of experimental conditions with a single or precomputed potential profile (e.g., Levitt, 1986; Cooper, Gates & Eisenberg, 1988a,b; Chiu & Jakobsson, 1989; Barcilon et al., 1993; Eisenberg, Klosek & Schuss, 1995).

These comments concerning consistency apply to theories which use rate constants to describe open channel permeation or gating (loc. cit.) even if those theories do not explicitly contain potential profiles, because rate constants always depend on the underlying potential profile. Even in the simplest most approximate theories of gas phase kinetics—like Eyring rate theory—rate constants are exponential functions of barrier height, i.e., the maximum of the potential profile $\Phi(\zeta)$ (Wigner, 1938; Laidler, 1969; Johnson, Eyring & Stover, 1974; Skinner & Wolyens, 1980; Truhlar & Garrett, 1984; Truhlar, Isaacson & Garrett, 1985). In more general theories—which provide the basis for rate theories in condensed phases and from which they must be derived (Chandler, 1978; Hynes, 1985; 1986; Berne, Borkovec & Straub, 1988; Fleming & Hänggi, 1993)—the rate constant also depends exponentially on the potential profile $\Phi(\zeta)$ (cf. Eisenberg, Klosek & Schuss, 1995, and Eq. (A9) and (A10) of the Appendix). In either case, anything that changes the potential profile will change the rate constant, often dramatically. Thus, rate constants of traditional theories—whether of gating, permeation, or 'active' transport—are likely to vary with experimental conditions, just as potential profiles vary.\footnote{If the permeation pathway is occluded in a closed channel, one would expect permeant ions to have much less effect on the opening process than on open-channel permeation, or on the closing process, for that matter.}

A protein or channel assumed to have an unchanging electric field is unrealistic because it almost certainly will violate the equations of the electric field; it is also likely to be a dead protein, unable to respond to its environment, the way live proteins do. In my opinion, biological control mechanisms are likely to modulate the electric field in channels and proteins, thereby initiating and governing conformational changes, as well as driving 'catalysis', analogous to permeation through the open channel (Moczydlowski, 1986; Eisenberg, 1990; Andersen & Koeppke, 1992).

### Sources of Diffusion

Permeation and diffusion are driven by gradients in concentration, which are the sources of flux. These concentrations are the sources of mass and free energy, if we use language of thermodynamics; they are the sources of the concentration and potential fields, if we use the language of the 19th century physics; they are the source of trajectories, if we use the language of probability theory and stochastic processes.

In biological membranes and channels, the concentration gradients arise from ions in the baths adjoining the membrane. The concentrations of these ions are maintained by ancillary experimental or biological systems that supply the ions equivalent to those that move through the channel and so sustain the free energy of the baths.

The concentrations in the baths are, in fact, usually taken as the only sources for diffusion; ions are not supplied within the channel, nor can they appear there spontaneously. Semiconductors are a little different because of the recombination process, but this is usually ignored in theories. Recombination does not occur in channels or solutions containing only strong electrolytes like Na$^+$, K$^+$, Ca$^{2+}$, or Cl$^-$. It can occur in other situations, see p. 10.

### Sources of the Electric Field

The sources for the electric field that drive the drift (i.e., migration) of ions are more complex than the sources for diffusion, whether in semiconductors, solutions, or channels. The sources are the several kinds of electrical charge in the system, each type with its own properties.

For example, only the charge in the baths, on the boundary of the system, connected to amplifiers, pulse generators or batteries, maintains the trans-membrane potential. That charge must be maintained by a continuous supply of energy from the outside world because flux flows across the channel dissipating energy and producing heat. Flux allows states (i.e., combinations of $q(x)$ and $C(x)$) that otherwise cannot occur. Thus, a theory without flux (e.g., PB$_n$) may miss the most interesting states and behaviors of a channel as it would miss most of the interesting states and behavior of transistors (which do not do much if they must start from a zero current condition and remain in or near to equilibrium: the amplification and switching which are the most important functions of transistors cannot occur at or near equilibrium).

The other types of charge (described below) also help create the electric field. They, however, cannot supply energy in the steady-state because they are not on a boundary of the system and so are not connected to an energy source outside the system.

---

\footnote{except under the conditions when constant field theory is a decent approximation (Chen & Eisenberg, 1991, and references therein).}
In free solution, the dominant charge is usually the mobile charge (i.e., ions) supplied at the boundaries of the system (i.e., at the electrodes) by the experimental apparatus that maintains the electrode potentials.

In channels or semiconductors, the dominant charge is usually the permanent charge because it has such a high density, several molar if described as a volume charge density, even if formal charges are absent (see pp. 8 and 17). The qualitative properties of the electric field in channels (and semiconductors), as well as its quantitative current voltage (IV) relations, are determined by the distribution of permanent charge, with a strong assist from the mobile charge (the ions) supporting the trans-membrane potential (analogous to the bias potential of a two terminal transistor).

Charge creates the electric field according to Coulomb’s law (if the charge is discrete) or Poisson’s equation (if the charge is distributed). The equations of the classical electric field describe the mapping between charge and potential both in the macroscopic world and in quantum chemistry (Hellman-Feynman theorem: Feynman, 1939; Deb, 1981; Mehra, 1994, p. 71; cf. Bader, 1990; Parr & Yang, 1989). All charge must be included in whichever equation(s) apply, because the effects of the dominant charge are significantly modified by the other charges. Linking potential and charge requires simultaneous solution of equations which describe how the electric field is created by charge (e.g., Poisson’s equation) and how charge flows and diffuses (e.g., the Nernst-Planck equations), and how energy is supplied to the system (boundary conditions). Together, these equations automatically describe the interactions of the different types of charge and the fields they create. Traditional theories of physics uses the Poisson and Nernst-Planck equations for this purpose, describing the mean concentrations and potentials (i.e., fields) of classical physics, but other theories are possible, and probably needed to describe the detailed properties (e.g., nonideal ratios of unidirectional fluxes of tracers, Jacquez, 1985; Hille, 1989; Chen & Eisenberg, 1993b) of correlated motions of random trajectories in single filing systems like channels.

### Poisson’s Equation in Semiconductors

Poisson’s equation has been used for some 40 years by solid-state physicists to describe the mean electrical field in semiconductors (Roosbroeck, 1950) and Shockley, (1950). Coupled to the Nernst-Planck equations, Poisson’s equation describes the many semiconductor devices with a wide range of characteristics. These equations are coupled to each other because the ions that create the electric field (as described by Poisson’s equation) flow and so are modified by the electric field (as described by the Nernst-Planck equations).

The combined Poisson and the Nernst-Planck equations have been derived and are used in several different experimental and theoretical traditions in physics and chemistry (Ashcroft & Mermin, 1976; Mason & McDaniel, 1988; Spohn, 1991; Balian, 1992; Jerome, 1995; at equilibrium see references to PB, most importantly, Davis & McCammon, 1990; Hecht et al., 1995; also, Honig & Nicholls, 1995), where they have been tested explicitly many times by simulation and experiment. The drift diffusion equations are tested implicitly every time we use a semiconductor device designed with them. There is little question in these fields that the equations are consistent with the laws of mathematics and physics, and (in particular) that they handle self-energy self-consistently in the mean field approximation 12; for example, energetics are discussed at length in the literature of PB, where self-energy is handled as it is in PNP at equilibrium (e.g., MacKerell, Sommer & Karplus, 1995; Antosiewicz, McCammon & Gilson, 1994; Gilson et al., 1993; Sharp & Honig, 1990; Jayaram et al., 1989; Gilson & Honig, 1988). The meaning of energy in a dissipative, nonlinear, nonequilibrium, and open system like PNP is discussed later in this paper, on p. 11.

We will discuss the validity of PNP as a description of open ionic channels in some detail later, starting on p. 13. Suffice it to say here that the PNP equations are a mean field theory—probably the simplest self-consistent mean field theory—that has proven helpful in many other fields, even where they are only approximately valid and the microscopic meaning of their parameters is not fully understood. The theory has been widely used with some success in many fields (cf. footnote 26). The most serious problem with the theory is usually the meaning of its parameters, that is to say, the relation of its effective parameters to the underlying parameters of atoms and their stochastic motion.

### Semiconductors, Solutions, and Channels

The analogy between current flow in ionic channels and solutions and other physical systems has been obvious for a long time but the analogy has not been very productive because most physical systems (e.g., semiconductors) contain permanent charge (doping) and ionic solutions do not. A (particular) distribution of permanent charge is what turns a semiconductor into a transistor, for example, and ionic solutions cannot be transistors because they do not have permanent charge. But channels have a large density of surface charge lining their...
tiny pore and so the analogy between semiconductors and channels is likely to be useful.

Permanent charge is more widely distributed and concentrated in proteins than is sometimes realized (see p. 17). Permanent charge is found in most atoms and bonds of proteins, not just in atoms with formal charges; many atoms of proteins contain between 0.1 and 0.6 elementary charges, for example, the carbon, nitrogen, oxygen, and perhaps even hydrogen in the amide bonds that link every amino-acid in a protein (Schultz & Schirmer, 1979; Fersht, 1985; McCammon & Harvey, 1987; Brooks, Karplus & Pettit, 1988; Creighton, 1993). The charge on each of these atoms is nearly as significant as a formal charge on a carbonyl or amine group of the protein, or, for that matter, on a permeating ion like Na\(^+\) or Cl\(^-\). The charge of each of these atoms contributes significantly to the electric field in and around channels and other proteins.

**Electrical Forces in Chemistry**

Electric fields are so important for channels and proteins because electrical forces dominate chemical phenomena. Indeed, in a certain sense, all of chemistry arises from electric charge and its interactions according to Coulomb’s law:

“... all forces on atomic nuclei\(^{16}\) in a molecule can be considered as purely classical attractions involving Coulomb’s law. The electron cloud distribution is prevented from collapsing by obeying Schrödinger’s equation.” (Feynman, 1939).

For this reason, any description of a protein must describe its electron cloud and nuclei, which form the distribution of electrical charge in the protein: its permanent charge. Any description of the movement of ions through a channel protein should keep track of all the charges in and around the channel for the same reason: the electrical charge determines the forces on the permeating ions, as well as the forces that make the protein move, the forces that drive its conformational changes.

**Charge in Channels**

The electric field in a channel arises from charge. PNP theory includes four types of charge:

\(^{(i)}\) the permanent charge on the channel protein that arises from its chemical structure. Permanent charge is often called bound charge (Griffiths, 1981; Purcell, 1985, Ch. 10) but that phrase is unfortunate because mobile charge is also bound to proteins, in a very real sense. Permanent charge is sometimes called fixed charge, but that phrase is also unfortunate (at least on the atomic scale) because the permanent charge is highly flexible, see p. 14.

\(^{(ii)}\) the contents of the channel, namely, the charge in the channel’s pore carried by the (average) concentration of ions there. The concentration of ions in the channel’s pore both determines and is determined by the electric field.

The double role of charge, as cause and effect of the field, can easily cause difficulties. That is exactly why the electric field was so difficult for 19th century physicists to understand, according to several historical accounts (Buchwald, 1985; Hunt, 1991; Siegel, 1991). Indeed, until the discovery of the electron, at the end of that century, the distinctions between permanent and induced charge were not understood. The electric field is, however, no more complicated than gravitational interactions of deformable (cf. ‘polarizable’) objects of similar mass, like binary stars. There, too, the location and shape of the sources are both determined by and determine the gravitational force.

Gravitation is not usually considered in this light (except by astronomers) because we usually compute gravitational forces between rigid objects of very different mass. Then, only one object determines the field, to a good approximation. The essential fact about electricity is that this approximation rarely applies: most objects, e.g., atoms and ions, have about the same amount of charge and many can move and deform (‘polarize’) in the electric field.

\(^{(iii)}\) the dielectric charge (that we call induced or polarization charge: Griffiths, 1989, Ch. 4; Purcell, 1985, Ch. 10) which is traditionally described by the volume density of (hypothetical) dipoles (of infinitesimal size) that represent the small movements of electrons and nuclei (etc.) induced by the electric field. Dielectric or polarization charge is defined as the induced charge movement (dipole moment per unit volume, to be precise) proportional to the local electric field.

Induced charge with more complex properties is usually described as a component of mobile charge that varies with the electric field. The ions inside a channel’s pore form a concentration of mobile charge that varies nonlinearly with the electric field.

Polarization or induced charge does not include dipoles independent of the electric field, so it does not include the macroscopic dipoles of say a carbonyl group, even though carbonyls are invariably called a ‘polar’ group! In our system of definitions, that kind of charge is described as part of the permanent charge.
(iv) boundary charge, namely, charge applied at the boundaries of the system during the experiment (by electronic and chemical apparatus) to keep the potential and concentrations (nearly) constant at those places.

Everyone who does experiments knows how important it is to measure and/or control the concentrations and potentials on both sides of channels that together determine the thermodynamic driving force, the gradient of electrochemical potential. These concentration and potentials supply the energy, charge, and matter that make currents and fluxes flow. They are the sources of free energy for channel phenomena and so, if they vary, nearly everything measured also varies.

For precisely the same reasons, theory must describe these sources carefully. In the present case (and usually) the sources are on the boundaries of the system, and are given by boundary conditions, when the system is described in three dimensions.

**Shielding of ‘Surface’ Charge**

Permanent charge on the surface of membranes (or at the ends of channels) attracts ions of opposite sign from the (overall) electrically neutral bathing solution, and thereby creates an ionic atmosphere with net charge (analogous to the ionic atmosphere of Debye-Hückel theory, here called the diffuse double layer or Gouy-Chapman layer (see extensive references in the channel context in Bockris and Reddy (1970), Bard and Faulkner (1980), McLaughlin (1989) and Green and Andersen (1991), are gateways to the electrochemistry literature). This charge produces a potential drop in the surrounding bathing solution, called the surface potential in physical chemistry, and the built-in potential in semiconductor physics. If the concentration of ions in the bath is very high, compared to the density of surface charge, the ionic atmosphere ‘shields’ or ‘screens’ the surface charge, and the surface or Donnan potential extends only a short distance into the bath. If the concentration of ions is low, the surface charge is hardly shielded, and the Donnan potential extends far into the solution.

In semiconductors, screening has long been known to be important. That literature of thousands of papers can be reached through Selberherr (1984) Rubinstein (1990) and Jerome (1995). In electrochemistry, screening has not been analyzed as carefully, usually because models have been reduced to avoid mathematical difficulties by ignoring one of the types of charge or another (cf. p. 8), or by setting flux to zero (e.g., de Levie & Moriera, 1972; de Levie, Seidah & Moreira, 1972; de Levie & Seidah, 1974; de Levie, Seidah & Moreira, 1974; Brumleve & Buck, 1978; Mafé, Pellicer & Aguillela, 1986, 1988; Mafé, Manzanares & Pellicer, 1988, 1990; Murphy, Manzanares, Mafé & Reiss, 1992; Nahir & Buck, 1993; Guiraro, Mafé, Manzanares & Ibáñez, 1995, who used the name Poisson-Nernst-Planck).

In channology, the effect of surface charge on open channel permeation has received much attention, see Apell et al., 1977; McLaughlin et al., 1981; McLaughlin et al., 1983; Dani, 1986; Green, Weiss & Andersen, 1987; Jordan, 1987; Kell & DeFelice, 1988; Peskoff & Bers, 1988; Jordan et al., 1989; McLaughlin, 1989; Cai & Jordan, 1990; Mathias et al., 1991; Green & Andersen, 1991.

In channology, the effect of surface charge on gating phenomena has been studied for many years (at least since Frankenhaeuser & Hodgkin, 1957; Chandler, Hodgkin & Maves, 1965; Gilbert & Ehrenstein, 1984, is a review; see also Hille, 1992). Divalent ions and pH also have large effects on gating. Both have been explained by their effect on the surface potential (McLaughlin et al., 1981; McLaughlin, 1989; Green & Andersen, 1991), although there is a huge literature investigating other more specific effects of the divalents (Hille, 1992, is a gateway to that literature).

In channology, the effect of shielding of the protein charge in the channel’s pore has received less attention, although some reduced models have been developed (e.g., Teorell, 1953; see also Mauro, 1962; Bruner, 1965a,b, 1967; Åqvist & Warshel, 1989; Edmonds, 1994) and Peng et al., 1992, and Zambrowicz and Colombini, 1993, apply Teorell’s theory (in essence) to modern data.

In channology (or anywhere else), what has been very difficult to develop has been a computable treatment of shielding that includes flux and allows coupling of the effects of all types of charge.

Only recently have computers been fast enough, and numerical techniques sophisticated enough to analyze the entire system, involving as it does coupled nonlinear partial differential equations describing flux in the presence of surface, induced, and fixed (i.e., permanent) charge (first successfully treated, in semiconductors, Fatemi, Jerome & Osher, 1991; later, independently, in channels, by Chen, cf. Chen & Eisenberg, 1993b). Our results using these new methods reinforce what others have suggested (Green and Andersen, 1991): surface charge has global effects; all types of charge in the channel interact with each other. Thus, understanding and predicting the effects of bath concentration and transmembrane potential requires the simultaneous solution of a system of nonlinear differential equations (Poisson and Nernst-Planck equations) and boundary conditions, including all types of charge.

It would be very helpful to have a reduced version of PNP which predicted the qualitative properties of currents and fluxes as experimental conditions or protein structure as changes. An equilibrium model (e.g., a version of PBE) does not seem helpful: comparison of equi-
librium and nonequilibrium situations in one dimension (Chen et al., 1995) have shown that the equilibrium distribution of potential is a surprisingly poor approximation to the nonequilibrium situation. We have tried to find a simple way to predict the nonequilibrium effects of PNP without solving the entire problem by manipulating the physical components of $\phi(x)$. For example we have subtracted surface potentials, polarization effects, potential profiles at zero current, and so on seeking an invariant curve—a single (reduced) profile of potential that would characterize the channel protein itself (independent of the subtracted effects). We have not succeeded: the reduced profiles we have evaluated have been as sensitive to concentration and trans-membrane potential as the original $\phi(x)$. Physicists and engineers have also been unsuccessful in constructing a reduced version of their drift-diffusion equations, despite strong incentives (e.g., the qualitative design of new transistors and many decades of trying (Coughran et al., 1994)). We and others have not given up this quest, but for now we must be content with the rapid, accurate, and convenient numerical solution of the model. The solution is easy to compute, and the program to compute it is available on request, but it must be recomputed for each experimental condition or each distribution of permanent charge. We do not know how to make qualitative predictions of the effects of change in permanent charge (i.e., protein structure), bath concentrations, or trans-membrane potential.

Changes in Permanent Charge

In most situations the distribution of permanent charge on a protein does not change. The exceptions, however, are of considerable biological importance. When the conformation of a channel protein changes for whatever reason, its distribution of permanent charge will change. When transmitters or second messengers bind to a protein (or phosphorylate it), they add or subtract charge and are likely to induce a conformational change. Both will change the distribution of permanent charge significantly and thereby change the electric field even if they are formally neutral, because nearly all transmitters and messengers are zwitterions or polar (otherwise, they would not be very soluble in water!). Indeed, a conformation change in any membrane protein will almost certainly modulate the electric field in the nearby lipid membrane and adjacent channels, thereby initiating and governing gating and open channel permeation.

When the chemical nature of the protein changes because of changes in covalent bonds, the distribution of permanent charge will also change. Obvious examples are phosphorylation (that produces local concentrations of $P_j = [H_2PO_4^-] + [HPO_4^{2-}]$) and changes in ionization of the acidic or basic groups in the protein (that produce local concentrations of $H_2O^+$ or $OH^-$) caused by changes in $pH$. If phosphorylation (or ionization, see next paragraph) actually occurs within a channel’s pore, it will change the concentration and species of current carriers there, as well as permanent charge. In that case, the dragging of ions through the pore by $P_j$, $H_2O^+$, or $OH^-$ (and vice versa) might play an important role in active transport (Hille, 1989; Läuger, 1991) or some types of gating (Hille, 1992).

Field Dependent Ionization

Experiments show (Tanford & Kirkwood, 1957; Edsall & Wyman, 1958) that the electric field can significantly change the ionization (i.e., permanent charge) of proteins by changing the effective $pK_a$ of hydrolyzable (often called ‘ionizable’) groups—Glu, Asp, Arg, Lys, His. The effects found experimentally have been extensively analyzed by the mean field Poisson-Boltzmann (PB) theory (Straatsma & McCammon, 1991; Yang et al., 1992; Antosiewicz, McCammon & Gilson, 1994; Potter, Gilson & McCammon, 1994; Rajasekaran, Jayaram & Honig, 1994; MacKerell, Sommer & Karplus, 1995). Of course, all the energy terms in PB or PNP contribute to the change in $pK_a$. In channel proteins, these include the energy of the membrane potential (i.e., of the charges on the boundary the membrane and in the bath), of the ions in the channel, of the permanent charge, of the induced charge, and of mechanical and electrical forces in the boundary conditions themselves (e.g., the ‘dielectric pressure’).

The possibility that permanent charge in a channel’s pore is modulated by membrane potential or phosphorylation (etc.) is tantalizing because such modulation might provide a link between the physics of the open channel and the elaborate biochemical mechanisms that control channel function (Gilbert & Ehrenstein, 1970; Cramer et al., 1995, p. 628). Ionizable residues are found in putative pore forming regions of many channels (review: Perachia, 1994) and mutations in them often have profound effects on channel function and modulation, both on gating and on open channel permeation (Miller, 1989; Sigworth, 1994; Montal, 1995, and references therein). If significant ionizable charge is present within a channel’s pore, complex interactions will occur: charge creates the field, the field modifies the charge, and the potential barriers to permeation will surely change in an interesting way, at least in PNP theory, and probably in the real channel as well, both transiently and at steady-state.\footnote{It should be mentioned that ‘activation curves’ showing the voltage dependence of many rate constants of (macroscopic) gating (Hodgkin & Huxley, 1952; Hille, 1992) have the shape of titration curves or their derivatives with respect to potential.}
Energy in Electrodiffusion Systems

The idea of energy is considerably more subtle than it sometimes seems, if systems are dissipative, nonlinear, and open, let alone chaotic, thus violating, in many respects, the assumptions made in elementary treatments of thermodynamics or mechanics. Explicit analysis of the mechanics of nonlinear mechanical systems containing just a few components, like a driven dissipative pendulum (Ben-Jacob et al., 1982; Kupferman et al., 1992), shows the need for careful thought and mathematical treatment before extending the laws of thermodynamics to nonequilibrium nonlinear systems.

Usual treatments of energy in mechanical systems require the boundary conditions of the system to be ‘holonomic’. These treatments can be extended to frictional systems if Rayleigh’s dissipation function can be defined (Goldstein, 1980, p. 21–25, p. 62). Most treatments of irreversible thermodynamics depend on the definition of such a function (e.g., Katchalsky & Curran, 1965). Treatments of energy in an equilibrium, quasi-thermodynamic theory like PB, (loc. cit., and also Gilson & Honig, 1988; Jayaram et al., 1989; Sharp & Honig, 1990; Gilson et al., 1993; Antosiewicz, McCammon & Gilson, 1994) cannot themselves define Rayleigh’s function because those theories assume equilibrium, and at equilibrium, atoms and charges have zero (mean) velocity, velocity dependent potentials do not appear (Goldstein, loc. cit.), no flux flows, and so frictional forces and dissipation are zero, as well.

It is not clear a priori whether Rayleigh’s dissipation function can be defined for a general electrodiffusion system because diffusion is strongly coupled to the electric field, and the system (with coupling) is inherently dissipative, nonlinear, and open, with flux of energy and matter (and charge) across its boundaries, as we have mentioned many times. The work needed to move charge in a protein or channel (e.g., to ionize residues in a channel) may or may not be describable by a path-independent energy function (even in principle: Griffiths, 1989, p. 187). It remains to be seen.

Whether or not energy and dissipation functions can be defined for these systems—and I surely hope they can!—any theory using energy as a variable must keep careful track of all its components.

---

18 As Goldstein puts it in the second edition of his classic text (p. 64, op. cit.): “A good deal has been written about Hamilton’s principle for nonholonomic systems, and most of it is wrong (including some things that were said in the first edition) . . . . [Rund, 1966, concludes that] Hamilton’s principle is applicable only to holonomic systems.”

19 which for present purposes, can be crudely read as ‘isolated’ although there is much more to the idea than that (Goldstein, 1980, p. 11, also see items listed under ‘constraints’ in his index).
All Charge Produces the Electric Field

The electric field, that determines the energy, arises from charge according to Poisson’s equation, which itself is just a restatement of Coulomb’s law.

To describe a channel or protein, either Poisson’s equation or Coulomb’s law needs to include as its sources all the types of charge described on p. 8 because each helps produce the field. The equations need to describe how the charge and potential at the boundaries of the baths and the boundaries of the channel (at its walls) change as experimental conditions change, e.g., as concentrations or potentials are changed in the bath: the sources (boundary conditions) in the theory must, of course, have the same properties as the charge in the physical world.

Interestingly, traditional theories of ionic solutions (see references in Bockris & Reddy, 1970; Bard & Faulkner, 1980) and liquids (see references in Allen & Tildesley, 1990), and simulations of molecular dynamics (Evans & Morris, 1990; Hoover, 1991; Mareschal & Holian, 1992; Lowe, Frenkel & Masters, 1995), even of channels (reviewed in Roux & Karplus, 1994), often pay little attention to charge on the boundaries20 and the resulting macroscopic electric fields, and so cannot deal with many phenomena of channels, membranes, cells (or tissues, for that matter), which are produced by charge at the boundary of the protein or bathing solutions. These theories cannot describe the ‘far fields’ or general boundary conditions that produce resting or electrotonic potentials (Jack, Noble & Tsien, 1975) of nerve and muscle cells, or syncytial tissues (Eisenberg & Mathias, 1980) or long insulated conductors bathed in an ocean of salt water, for that matter (i.e., the Atlantic cable: Kelvin, 1855, 1856). They cannot describe the qualitative mechanism of the action potential (Hodgkin & Huxley, 1952; Hodgkin, 1971, 1992) because that involves feedback between the membrane potential (i.e., boundary potential) and the potential within and current flowing through a single channel molecule (Hodgkin, 1958). Much recent work on liquids, ionic solutions, and proteins has been devoted to dealing correctly with just the Dirichlet far field condition of (uniform) zero electrical potential at infinity (using Ewald sums or reactions fields, viz.:

Solving the PNP Equations Means Computing the Electric Field

The PNP system of equations can be solved simultaneously, without approximation, using the numerical methods described in the Appendix. Even though PNP is a quite complex, coupled set of nonlinear differential equations, the underlying physical principles are simple: all types of charge in the channel, and at its ends, have interacting effects on the distribution of potential and concentration in the channel. The shape of the electric field is found to change substantially, even qualitatively, by some $k_B T e$ in most locations, in hundreds of different calculations simulating real experimental conditions, like those used in the laboratory, in which concentrations of ions are changed in the baths. Modulation of the electric field in channels is the rule, not the exception. As experimental conditions change, the potential profile changes.

Nonetheless, the surface of a channel or protein has often been described as a more or less unvarying set of potential barriers over which flux diffuses (e.g., Levitt, 1986; Cooper, Gates & Eisenberg, 1988a,b; Chiu & Jakobsson, 1989; Barcilon et al., 1993; Eisenberg, Klosek & Schuss, 1995) with a rate constant independent of the concentration of ions in the bath or channel. Hill (1977), Walsh (1979), Hill (1985), Hille (1992), Andersen and Koepppe (1992), Läuger (1991) and Eisenman and Horn (1983) are examples of this extensive literature of rate constant models. Such models of flux over barriers (whether using rate constants or not) implicitly describe the surface of a protein as an unvarying potential profile—a ‘potential of mean force’ in the technical language of statistical mechanics—thereby ignoring the effects of shielding, even though those effects can be large at the ends of the channel (loc. cit. and see Eqs. (A2)–(A4) of the Appendix) and in its pore (cf. Eqs. (A9) and (A10) of the Appendix) and are large in many other physical systems (as described on p. 9). Unfortunately, most theories of open channels are subject to these criticisms whether they describe permeation by rate constants or diffusion over barriers loc. cit. These theories do not allow the rate constants or barrier shapes to vary as concentration or trans membrane potential varies and
so they imply unrealistic properties of the electric field. In fact, if used over a range of experimental conditions, they are inconsistent with the laws showing how charge creates an electric field, namely Poisson’s equation or Coulomb’s law.

The Surface of a Protein is a Distribution of Charge, Not Potential

A channel (or a protein, for that matter) can certainly be described as a distribution of potential under one set of conditions. Indeed, it can be described as the same distribution of potential under another set of conditions, as long as all the charges (cf. p. 8) in the system stay the same. But if concentrations, or trans-membrane potential are changed (or if the protein binds a substrate or transmitter, or if it changes conformation), charges will change and so will the potential-of-mean-force.\footnote{If the conformation changes, the distribution of permanent charge will change as well, of course.}

After any charge in the system changes, a profile of the potential-of-mean-force can still be used to describe the channel protein, but it will be a different profile with a different size and shape. The potential-of-mean-force might conceivably stay the same at one location, but only if that location is connected to a battery that supplies or withdraws charge to compensate for changes in other types of charge. The potential profile can stay the same only if many locations are connected to different batteries, each of which supplies or withdraws the charge necessary to keep its potential constant.

Describing a protein as a potential surface is equivalent to describing that surface as a source of potential, i.e., as a Dirichlet boundary condition, a surface where the potential is fixed. A surface of matter cannot be a source of potential, described by a Dirichlet boundary condition, unless it is a metal connected to a battery. The surface of a protein is neither metallic, nor connected to multiple batteries and so the surface of a protein must be described as a distribution of permanent charge.\footnote{This is a first approximation; a second approximation, dielectric properties must also be included as they are in our analysis (cf. Eq. 1).}

Only the permanent charge stays constant as experimental conditions vary and change the mobile charge (i.e., ions) in a channel and at its boundaries. If those mobile charges vary, the total charge varies, and so does the potential. Thus, as experimental conditions are varied, as the potential or concentrations in the bath are varied, the electric field varies because the charge in the channel’s pore and at its boundaries vary, even if the (fixed) charge on the protein stays the same. The electric field is modulated by changes in experimental conditions: modulation is the rule, not the exception.

Because the flux through a channel is an exponential function of the potential profile $\Phi(z)$,\footnote{See Eq. (A8).} modulation of the electric field tends to govern the qualitative properties of channels. In this way, changes in the contents of the channel can govern the qualitative properties of current flow.

Validity of PNP Theory

Eisenberg, Klosek & Schuss, 1995 (following the simulations of Cooper, Jakobsson & Wolyne, 1985; Chiu & Jakobsson, 1989; and the lead of Barclon et al., 1993) proved the validity of the Nernst-Planck (NP) part of the theory for discrete atomic systems without single filing. E. Barkai, R.S. Eisenberg and Z. Schuss (1995) submitted, are extending the theory to single ion channels. Using only mathematics, Eisenberg, Klosek and Schuss (1995) showed that the NP equations describe the mean of the probability density function for location in systems dominated by friction, in which atoms follow the Langevin model of atomic motion, i.e., Newton’s laws plus fluctuations caused by atomic collisions.

The Langevin model is the starting place for most theories of atomic motion in condensed phases (Kramers, 1940; Gardiner, 1985). Even if the system is not entirely dominated by friction, and the full Langevin equation is needed, or the friction has complex properties, and the generalized Langevin equation is needed (Hynes, 1985, 1986; Berne, Borkovec & Straub, 1988; Hägggi, Talker & Borkovec, 1990; Fleming & Hägggi, 1993; Tuckerman & Berne, 1993), Eisenberg, Klosek and Schuss prove (using mathematics alone) that something very like Nernst-Planck is appropriate (their Eq. 6.15). Surprisingly, the main results of their analysis depend only on mathematical identities. All that is necessary for their proof is the existence of conditional probabilities of location. The conditional probabilities can be directly derived from simulations of molecular dynamics (without using even the generalized Langevin equation, e.g., see Barclon et al., 1993, Figs. 4 and 5) or from the Onsager-Machlup action formulation of Newton’s laws in the presence of thermal agitation (Onsager & Machlup, 1953; see modern application: Elber, 1996).

Thus, there seems little question of the validity of the NP equations for a stochastic discrete atomic system. They describe the mean value of the probability density function of location of ions within the open channel. The question then is how to compute the mean potential used in the NP equations.

The Mean Potential

The Poisson equation can describe a potential on many time scales and it can describe the mean potential pro-
duced by mean charge. In the PNP theory of channels, it describes the mean field approximation to the potential in the open channel produced by the average distribution of permanent, mobile, and induced charge (cf. Eq. (1)). In particular, it describes the potential averaged over the 5 μsec needed to measure a single digital sample of mean current (because of instrumentation and signal to noise problems). Five μsec is much slower than the time scales that determine the energetics of solvation (0.01 psec, Strait, 1995), or typical atomic motions in proteins (0.1 psec) and so averaging of these motions is necessary. Five μsec is also slow compared to the time scale of ion permeation: a univalent ion crossing a membrane every 160 nsec carries 1 μA of current and a single ion takes roughly the same time for its (first) passage across a channel, namely, its mean first-passage time (MFPT), estimated by Cooper, Jakobsson and Wolynes (1985); Chiu and Jakobsson (1989); Barcilon et al. (1993); Eisenberg, Klosek & Schuss (1995).

The time and length scale of atomic motions is known (Berry, Rice & Ross, 1980) from experiment (e.g., measurements of temperature factors of proteins by x-ray diffraction: Frauenfelder, Petsko & Tsernoglou, 1979; Parak & Knapp, 1984; McCammon & Harvey, 1987; Brooks, Karplus & Pettitt, 1988; Smith et al., 1990; Kuriyan et al., 1991); from general theoretical considerations (e.g., kinetic theory, McQuarrie, 1976; Gard, 1995); most vividly from the simulations of molecular dynamics in general (Burkert & Allgower, 1982; see Davidson, 1993, which introduces a review of the ‘state of the art’: Chem. Reviews, 93(7), 1993); from simulations of proteins (Brooks, Karplus & Pettitt, 1988; McCammon & Harvey, 1987); and from simulations of channel proteins, in particular gramicidin (reviewed in Roux & Karplus, 1994), including our own (Elber et al., 1995). The range of time scales in proteins is remarkably large: experiments show conformation changes ranging from $10^{-12}$ sec to $10^8$ sec (Brooks, Karplus & Pettitt, 1988: Table I, p. 19).

The rms deviations of atoms from their mean positions are significantly larger (op. cit.), and the motions more violent, than imagined in traditional theories of channel permeation (Hille, 1992; Andersen & Koeppke, 1992; cf. Purcell, 1977; Berg, 1983). For example, comparing a snapshot of a simulation of a gramicidin channel with one taken some 10 psec later shows that most atoms are more than 1 Å away from where they were initially. If the electrostatic potential is computed from each snapshot, striking variations are found. The potential at most locations has changed more than $k_BT/e$: Coulomb’s law shows that a 1 Å displacement of a carbonyl oxygen (containing ~0.6e of charge), produces a potential change $>1/k_BT/e$ at distances of a few Ångstroms (e.g., Israelachvili, 1985, Ch. 3). Thus, potential profiles in pores change substantially at least every 10 psec: they are likely to fluctuate (by several $k_BT/e$) many times indeed while one sample of open channel current is measured (~5 μsec).

It is the fluctuating potential that determines the current measured in channel experiments. The experimental measurement of current is taken over such a long time period (compared to atomic or ionic motions) that it is necessarily an average. Thus, the potential that determines the measured current must be an average too, of some sort or other. The functions being averaged vary wildly, containing some 5 μsec/(MFPT = $10^{-7}$) = 50 or 5 μsec/(Solvation time period = $10^{-15}$) = $5 \times 10^9$ fluctuations in the 5 μsec period of a single current measurement. The average cannot be determined by simply summing the potential determined in individual simulations because that summation procedure is numerically (inherently) ill-determined. Furthermore, the various averaging procedures of equilibrium statistical mechanics do not apply to these systems far from equilibrium, and attempts to use them have not been successful (Allen & Tildesley, 1990, Ch. 8–11; Evans & Morriss, 1990; Hoover, 1991; Haile, 1992; Ch. 8; Mareschal & Holian, 1992); indeed, problems with equilibrium statistical mechanics have led some authors to propose a radical reworking of the ‘thermodynamic’ theory of flux (see experiments of Keizer and Chang (1987) and Hjelmfelt and Ross (1995) and discussion and references in Keizer (1987a,b) and Vlad and Ross (1994a,b) and Peng et al. (1995)). Rather, the average must be determined separately, by its own theory, just the way the average of the trajectories of a Brownian motion are determined by Fokker-Planck equations (Arnold, 1974; Schuss, 1980; Gardiner, 1985; Gard, 1988), not by actually adding up the trajectories.

The nature of the average potential depends on the atomic properties of the channel, as it does in any macroscopic condensed phase. If the motions of the permeating ion (or more precisely the permion) are much slower than the atomic motions surrounding it, the Hartree self-consistent field (SCF) approximation (Ashcroft & Mermin, 1976; Kittel, 1976) is usually invoked: the effective potential for an ion is determined from the averaged locations of the other charges, i.e., atoms.

It is certainly possible to imagine or build a system

---

24 These rapid fluctuations probably account for the very large open channel noise we (and others) have observed at high frequencies. Below (say) 10 kHz—where open channel noise has been reported—these rapid large fluctuations presumably ‘average out’ to zero.

25 The slightest rounding error, error in truncation of the sum, or systematic error would dominate the estimate of such a sum.

26 Mathematically speaking, the trajectories are ‘functions of unbounded variation,’ oscillating an infinite number of times in any finite interval, no matter how small (Wong & Hajek, 1985, p. 53) and thus are hard to evaluate, represent, or approximate in a finite simulation or calculation.
where macroscopic flux does not depend solely on the averaged potential, but rather depends on some extremely weighted average or particular extreme value of the fluctuating atomic potential, and so the SCF approximation fails. But if macroscopic properties depend on such extraordinary events, they cannot be described literally by classical macroscopic theories of average properties like Maxwell’s equations and diffusion equations of Fick’s law, or statistical mechanics as built on the Boltzmann transport equation. Classical macroscopic theories of matter are literally valid only when the averages of atomic trajectories are well behaved.

Most theories of condensed phases, solid state, or gas phases assume that the mean dependent variables (like potential) obey the macroscopic laws of electrostatics, and the fluxes are described by macroscopic laws of diffusion of charged particles, i.e., the Nernst-Planck equations, even though the electrical potential and other parameters are known (by direct experimental measurement in many of these sciences) to fluctuate wildly on atomic scales of length and time. These theories27 all use macroscopic laws like Poisson’s equation and the NP equations to predict macroscopic measurements. They all use SCF and effective parameters to describe macroscopic properties as averages of atomic trajectories. And all these macroscopic theories fit a wide range of the phenomena seen in ordinary experiments—the ordinary phenomena of classical physics. Of course, they cannot describe the extraordinary phenomena of modern physics (e.g., conduction of current by holes in semiconductors), precisely because those phenomena depend on correlations that are ignored in the averages of traditional macroscopic theory.

PNP theory must be viewed then as a continuum theory of the mean field much like the other continuum theories of nineteenth century physics. It was the triumph of nineteenth century physics to realize that simple statements of seemingly vague conservation laws become powerful and specific predictors of complex experimental phenomena (e.g., water waves breaking on a shoreline) when converted to mathematical form (e.g., partial differential equations) and coupled to boundary conditions (to describe the source of energy and flow) and constitutive relations (e.g., Ohm’s law, Fick’s law) that describe the mean flux of uncorrelated particles.

PNP is perhaps the simplest theory that describes the diffusion and migration of ions in the presence of an electric field that also computes the field from the charges present; in that sense, it is the simplest mean field theory consistent with the laws relating charge and electrical potential. But it certainly is not the only such theory; others may in fact be much better descriptions of the correlated motion of ion, water, and atoms (of the channel protein) that accompany permeation (convincingly shown experimentally by Rosenberg & Finkelstein, 1978a,b; Levitt et al., 1978; Finkelstein & Andersen, 1981; Dani & Levitt, 1981; Levitt, 1984; Finkelstein, 1987; Hille, 1992; and shown in simulations (Chiu et al., 1989; Roux & Karplus, 1991; Roux & Karplus, 1994; Elber et al., 1995)).

### Quasi-Particles as Correlated Motions in a Mean Field

It is instructive to consider how the correlated motions of extraordinary phenomena are analyzed in other sciences (Ashcroft & Mermin, 1976; Mason & McDaniel, 1988; Cox, 1993) where direct measurements of atomic and transport properties are routine, along with extensive simulations (Bird, 1994). In these sciences, mean-field theories are common, despite the universality of correlated motions, and mean field theories like PNP (or identical to PNP: Ashcroft & Mermin, 1976; Ch. 28 and 29) are used successfully to fit a wide range of qualitative and quantitative behavior (Mahan, 1993). The PNP equations appear because they arise from conservation laws (that are true no matter what are the details of the atomic motion) and simple constitutive laws, like Fick’s law, which approximately describe a wide range of systems. But these mean field theories are often not literally true as descriptions of motions of individual atoms, because the atomic motions are extraordinarily correlated. Rather, the mean field theories are used28 to describe collective motions as if they arose from the rigid body translation of a group of atoms (Goldstein, 1980: Ch. 5). The (sometimes hypothetical) rigid body is called a ‘quasi-particle.’29 One of the main goals in the study of such physical systems is to establish the existence of such quasi-particles (by experiment, simulation, and theory) and to determine their conservation laws, laws of motion, and effective parameters.

Consider, for example, the extraordinary phenomena of ferromagnetism, superconductivity, and polarons (Ch.

---


28 A mathematician might say “they are used figuratively” when speaking of those cases where the quasi-particle and its behavior cannot be rigorously derived from the underlying atomic dynamics.

29 The discussion of quasi-particles in Ziman (1960), Smith and Jensen (1989), Ch. 5, and Ashcroft and Mermin (1976) were particularly helpful to me.
find a range of phenomena (e.g., of the open channel) that are well described by the mean field theory presented in the Appendix, but with effective parameters whose meaning can be difficult to sort out, because the object moving in the mean field theory is not an atom or molecule but a quasi-particle we call ‘a permion’ (Elber et al., 1995), that moves along an effective reaction coordinate, a twisted path, that is usually not perpendicular to the membrane surface. Indeed, classical phenomena of open channel permeation arise from the correlated motion of the permeating ion and water, as transport experiments convincingly showed some time ago: (Rosenberg & Finkelstein, 1978a,b; Levitt et al., 1978; Finkelstein & Andersen, 1981; Dani & Levitt, 1981; Levitt, 1984; Finkelstein, 1987; Hille, 1992). The correlated motions are, of course, also apparent in simulations (Chiu et al., 1989; Roux & Karplus, 1991; Roux & Karplus, 1994; Elber et al., 1995). Other classical phenomena depend on the interactions of ions coming from different sides of the membrane (e.g., nonideal ratios of unidirectional fluxes of tracers, Jacquez, 1985; Hille, 1989; Chen & Eisenberg, 1993b) and so must be described by a theory that allows such interactions, for example, a mean field theory of the correlated diffusion of a permion in a single file (E. Barkai et al., 1995 submitted). Of course, the utility of this idea of a permion remains to be established. Now, it is a compelling and not very new image; eventually, it may become a full fledged theory predicting selectivity, fluctuations, flux, and blocking phenomena in open channels (Hille, 1992).

It is even possible that the phenomena we call (single channel) gating will be best described as the motion of a quasi-particle (perhaps, the permion; more likely a quasi-particle with different properties, a ‘gate-on’) that follows its own laws of motion (i.e., Langevin equation) with probability density described by something like PNP. If that motion occurs over a high barrier, its mean time course will be exponential and rate theory will be a good approximation, as it is to most phenomena of single channel gating (McManus & Magleby, 1988; McManus et al., 1988; McManus & Magleby, 1991).

PNP is thus an appropriate mean field theory of the rapidly fluctuating atomic scale potentials of the open channel, as long as averages are taken on the time scale at which current is measured (e.g., over 5\,\mu s, which is long compared to the permeation time of a single ion), and as long as the system is reasonably homogenous for that 5\,\mu s.

**Homogeneity of the Open State**

The homogeneity of states of proteins cannot be assumed (Frauenfelder, 1985; Ansari et al., 1985; Frauenfelder, Sligar & Wolynes, 1991) and so it is fortunate that the homogeneity of the open state of a channel is known.
directly from experimental measurements of single channel currents on the biological time scale. Measurements of open channel noise (chiefly from Sigworth’s lab, starting with Sigworth, 1985; for more recent references see Heinemann & Sigworth, 1991, also Hainsworth, Levis & Eisenberg, 1995) show little correlation between successive samples of the current records; currents hardly vary within one channel opening or from opening to opening, or even experiment to experiment (when recording conditions are the same). The variance and power spectrum within a single prolonged opening is the same as the variance from opening to opening (op. cit.). Because these measurements are made on the actual time scale of biological function, they estimate the inhomogeneity relevant for that biological function. In fact, they show directly that the open channel is quite homogeneous on the functionally relevant time scale. Of course, they say nothing about much faster time scales; inhomogeneity of the type seen by Frauenfelder (1985), Ansari et al. (1985); Frauenfelder, Slijer and Woyles (1991) might well exist on the picosecond time scale while the channel appeared homogeneous on the μsec time scale of permeation.

The homogeneity of currents observed experimentally (on the functionally relevant time scale) implies an underlying structural homogeneity of the open state of channel proteins, on the same time scale: that is the time scale in which the mean quantities of PNP theory are computed (namely, the parameters and variables of Eqs. (1) and (2)). The conformation and the shape of the electric field must be reasonably constant and nearly the same whenever the channel is open on the time scale of μsec; otherwise, the experimentally measured currents would not be reasonably constant and nearly the same on that time scale. 31

The question is what do the words ‘reasonable’ and ‘nearly the same’ mean? I believe this question can be answered directly from experimental data because the open channel noise itself is a measure of the inhomogeneity of the open state. Indeed, it is an upper bound on the inhomogeneity of the open state because open channel noise can arise in other ways besides inhomogeneity.

In fact, the measured current is a more sensitive measure of the homogeneity of the open state than the potential profile: current through a channel is an exponential function of the potential profile $\Phi(\zeta)$ in nearly all theories (e.g., as shown explicitly in Eq. (A8), left hand side, Appendix), and the concentration is also a steep function of potential as shown on the right hand side of Eq. (A7); consider special case with $C(J) = 0, V_{appl} = 0$. Thus, if single channel current is found to be reasonably constant experimentally, the contents of the channel are likely to be reasonably constant, and the potential profile is likely to be even more constant.

We should thus not be surprised that open channel noise is reasonably well behaved in most channels. It does not differ from instrumentation noise by more than a factor of 2× in the papers of Sigworth (where it is often less than that) or Hainsworth, or in most other channels. Indeed, if uncertainty in the level of open channel current is comparable to the mean amplitude of open channel current, or if the current fluctuates significantly (say 80% deviation > 10% of open channel current) on a time scale comparable to the (mean) duration of open channel current, the records do not fit within the paradigm of single channel recording (Bean et al., 1969; Hadley & Haydon, 1970; Neher & Sakmann, 1976; Ehrenstein & Lecar, 1977; Sigworth & Neher, 1980; Neher, 1982; Sakmann & Neher, 1983, 1995) and are likely neither to have been pursued (very far) nor to have been reported (in full length publications).

The homogeneity of the open state is particularly interesting, because states of proteins in general are not considered so homogeneous (op. cit). The homogeneity of the open state is probably a consequence of the time scale of permeation and the high density of fixed (i.e., permanent) charge along its wall. If that charge is referred to the volume of the channel’s pore, its concentration is several molar; e.g., in gramicidin, in which the carbonyl oxygen’s make up the wall of the channel, each 3 Å turn of the helix contains about 0.6e, giving a concentration of permanent charge of 2.6 Molar. Such a highly charged tunnel is likely to change potential dramatically if its contents change and leave substantial fixed charge unshielded on the μsec time scale. And it may be impossible for such a large potential to be sustained for the functionally relevant time of 5 μsec, without the channel returning to a closed state. It seems possible that an open state of a channel is a quite special state, the only state (within some tolerance, of course) in which the contents of the channel—and thus its potential profile—are compatible with substantial flux. In this view, the open state is a conditional state; any of the many conditions (e.g., thermal fluctuations in concentration or conformation) that might perturb its balance of mobile and permanent charge would likely interrupt current flow, thereafter removing the protein from the open state. A closed channel would then be characterized by many states (i.e., potential profiles), while an open channel would be characterized by just one (within a certain tolerance, of course). If sufficiently well defined, the open state might be viewed as an (electrostatic) eigenstate of the protein, forming the ionic wire postulated by Rosenbusch (1988).

---

31 The subconductance states and flickers commonly seen in wide bandwidth recordings of channels also set bounds on the validity of our argument: they are probably a measure of another kind of inhomogeneity of the open state, distinct from that observed in measurements of open channel noise.
PNP as a Mean Field Theory

We conclude that the potential of PNP theory is a well-defined mean potential appropriate to describe current flowing in open channels on time scales longer than some 5 µsec. PNP should be viewed as the first order SCF theory of flux in channels relating mean potentials and currents and concentrations using effective parameters appropriate for the dynamics of ions and channel atoms on the 5 µsec time scale.

PNP theory does for channels, in the presence of flux, what PBₙ theory has done so successfully for proteins in the absence of flux (for some time now). PNP is a practical theory because of the advances in numerical analysis that allow quick computation of the full coupled nonlinear system. As far as we know, PNP is the first channel theory to solve the full set of equations involving induced, surface, and fixed charge, and flux through the channel. It is certainly not the only such theory, nor can it be expected to be the best; one dealing more realistically with single filing would clearly be better.

Profile of Permanent Charge

The profile of permanent charge is perhaps the most interesting effective parameter of PNP theory, because it determines the qualitative properties of the open channel (as doping determines the qualitative properties of transistors, along with the bias voltages) and because it can be modified so easily nowadays by changes in the amino acid sequence of proteins once the gene for that protein is known and cloned. Of course, only the primary structure—the sequence of amino acids—can be read from the genome: although that sequence is thought to determine the three dimensional folding pattern of the polypeptide chain, only nature knows how to do that (Creighton, 1992): no one can predict three dimensional structure (indeed even refine a bad guess) at the present time. We have seen (on page 13) how sensitive the potential profile is to the location of charge, and so we can see how difficult it is to guess the biologically relevant profile of permanent charge without knowing the three dimensional structure, heroic attempts notwithstanding. Nonetheless, changes in the primary sequence, which may be assumed (with various degrees of certainty) to leave the rest of the three dimensional structure essentially unchanged, can often be interpreted (Miller, 1989; Perachia, 1994; Sigworth, 1994; Montal, 1995, and references therein). Indeed, some changes have strikingly specific results, presumably because they mimic a single amino acid substitution that evolution found useful, sometime ago.

Successful and complete investigation of structure function relations, of course, will need measurement of structure. Three-dimensional structures are much harder to resolve than primary sequences, but some channel structures are now known at atomic resolution (e.g., Cowan et al., 1992) and more will be. The question then is how do we determine the one dimensional distribution of permanent charge of PNP theory from the three dimensional structure of the channel protein (on the one hand) or the current voltage relations of the open channel (on the other)? Certainly, there should be a good qualitative agreement (i.e., location and number of maxima and minima) between the potential computed by three dimensional PBₙ theory and that used in one dimensional PNP theory (in the zero flux and zero membrane potential condition used in PBₙ theory). But the quantitative relation is hard to predict, because it will depend on the meaning of the effective parameters of PNP and PBₙ theory. That meaning depends on the structure and dynamics of the channel protein, on the µsec time scale of permeation, because the relation of charge and potential depends on the shielding (i.e., polarization) within the channel protein itself and the dominant source of shielding is the movement of the (charged) atoms of the protein. These are difficult problems theoretically and experimentally in every field, not just in proteins and PNP theory. Nor can they easily be solved by simulation. Effective parameters can rarely be derived (by mathematics alone) from models of atomic structure and motion in any system, including those known much better than proteins.

It is best now simply to ask these questions, waiting for another day, when perhaps answers will be available, at least for a specific channel. Perhaps, when we know the answers for a few specific channels, generalizations and conclusions concerning the atomic meaning of effective parameters will emerge.

General Rules

Our analysis implies general rules independent of the limitations of mean-field, or quasi-particle theories. I believe any theory must explicitly display the dependence of flux and current on the diffusion constant of the system: in liquids, solutions, and other condensed phases, nothing can be assumed independent of friction. I believe any theory must be consistent with the properties of the electric field and so

(1) the theory must compute the electric field, not assume it.
(2) the theory must compute the electric field from all types of charge (cf. p. 8).
(3) the theory must never assume that a potential is maintained constant at some location in matter (as experimental conditions are changed), unless that location is connected by a wire to experimental apparatus which
serves as a source of energy and charge (i.e., a battery or amplifier).

(4) the theory must recompute rate constants when experimental conditions change, anything that changes the electric field is almost certain to change rate constants. A protein or channel with an unchanging electric field is likely to be a dead protein, unresponsive to its environment, at least compared to natural proteins or channels.

(5) If the theory is to compute the current flow or flux on the experimental time scale of μsec (or longer), it must use a mean potential, averaged or computed (by some scheme or other) on the same time scale.

Conclusion

A theory that follows these rules will predict a live channel, a live wire whose electric field (flux, and thus rate constants) will vary with experimental and biological conditions. The rich behavior of transistors (and the computers made from them) arises from that variation of the electric field with conditions (i.e., boundary potentials). Perhaps, some of the oriented, coherent, and constructive behavior of channel proteins (that help form the ‘microscopic cybernetics’ of life) arises the same way, by variation of the shape of the field (and the resulting effects on channel permeation and gating, and thus transport). We must learn to compute the field if we are to understand how ‘it’ works, whether ‘it’ is a transistor, channel, or protein.

This paper describes the work of my collaborators as much as my own: it is a joy to have shared this adventure with Duan Chen and our extraordinary coworkers, Victor Barrilon, Ron Elber, Joe Jerome, Mark Ratner, and Zeev Schwartz. It is an even greater joy to thank them. My colleagues at Rush have contributed in more ways than I can enumerate. I am particularly grateful to Eduardo Rios for helping me interact with the classical traditions of the field and to him for his many other questions and suggestions (which stimulated, among other things, footnotes 15, 16, 24).

More specific thanks are due to Lou DeFelice, who suggested that I write this paper; to Joe Blum, who requested the Appendix; to Duan Chen, who suggested the metaphor of the dead channel locked into one occupancy state; to Olaf Andersen, who motivated me to relate PNP to traditional treatments of surface charge; to Tom Dussey, who suggested the ideas in footnote 12 and 31, among others; to Ron Elber, who first used the phrase characteristic (i.e., eigen-) state to describe the open channel; to Richard Henderson, who forced me to think clearly about fluctuations, their means and meaning; and to Zeev Schuss and Eberhard von Kitzing for illuminating discussions of the idea of energy.

I have no words that can express my gratitude for the steadfast support of Dr. Andrew Thomson and the National Science Foundation. Without that, this work could not have been done.

These words were written with BMV 540 ringing in my mind, thanks to Jill.

References


32 paraphrase of Monod, 1972, p. 45.


Appendix

PNP Theory

The historical antecedents to PNP and its relationship to PB theory (e.g., of shielding), which are close, have been discussed in the text (p. 9). Our contributions have been reported in a series of papers starting with Barcilon (1992). Barcilon considers the full three-dimensional problem and derives the appropriate form of the perturbation expansion. Incidentally, (in a result overlooked by many biophysicists) he solves analytically the electrostatic problem of a finite length cylinder embedded in a thin membrane, giving two different exact expressions for the potential, thus contributing importantly to the electrostatics problem posed first (I believe) by Parsegian. 1969. Barcilon et al. (1992) and Chen et al. (1991) derive and solve the one-dimensional theory without permanent charge. They show when a constant (electric field) or constant gradient (of concentration) can approximate the full equations (without permanent charge). Chen and Eisenberg (1993a) put permanent charge in the theory; Chen and Eisenberg (1993b) introduce nonequilibrium boundary conditions that allow a channel of one structure to produce single filing, and flux coupling reminiscent (in some ways) of mediated transporters. Interestingly, it has just come to our attention that the theory of bulk ternary ionic solutions includes flux coupling (Wendy, 1965) of positive or negative sign (Vitagliano & Santoro, 1970). Eisenberg (1996) embeds PNP theory in a hierarchy of models of different resolutions; Park et al. (1995) analyze the qualitative properties of the theory in the absence of permanent charge and built-in potentials. Chen et al. (1995) generalize PNP theory to include explicit equations for the consolidation and flow of heat, energy, and mass. Eisenberg, Klosek and Schuss (1996) derive (just) the Nernst-Planck equations from a stochastic analysis of flux over barriers began in Barcilon et al. (1993). Barkai, Eisenberg & Schuss (1995) extend the stochastic analysis to a one ion channel.

Because of this extensive documentation, here we only state the main equations used to predict the results of a typical experiment in which the single channel current $I$ is studied as a function of the trans-membrane potential applied to the baths $V_{appl}$ in a variety of solutions of different composition $C_i(0)$ and $C_i(d)$. The theory used to predict this current starts with Poisson’s equation (written here in dimensional form), which determines the potential $\Phi$ (units: volts) from the charges present (cf. p. 8)

$$-\varepsilon_{i,\sigma} \frac{d^2 \Phi}{dx^2} = e P(x) + e \sum_i \varepsilon_i C_i(x)$$

Induced Charge

Channel Contents

Permanent Charge

Deviation From Constant Field

$$\Delta = \Phi_i(0) - \Phi_i(d) = \frac{k_B T}{e} V_{appl}$$

Note that the potential $\Delta$ is not the trans-membrane potential $V_{appl}$ applied to the baths. The baths are assumed at equilibrium, even when current flows, so

$$C_i(0) = C_i(0) \exp[-\varepsilon_i \Phi_i(0)]$$

$$C_i(d) = C_i(R) \exp[-\varepsilon_i \Phi_i(d)]$$

Nernst-Planck equations determine the flux $J_i$ of each ion

$$J_i = -D_i \left[ \frac{d C_i}{dx} + \frac{d \Phi}{dx} \right]$$

where $D_i$ is its diffusion constant.
The Nernst-Planck equation (A6) can be integrated analytically to give expressions for the concentration of ions in the channel, namely the channel's contents

\[
C_{j,x} = \frac{C_j(L) \exp(-\gamma z_j V_{applied}) + \Phi \Phi_j \varepsilon_j}{\int_0^L \exp \left( -\frac{\gamma z_j \Phi \Phi_j \varepsilon_j}{L} \right) dx}
\]

\[
+ \frac{C_j(R) \exp(-\gamma z_j V_{applied}) - \Phi \Phi_j \varepsilon_j}{\int_0^L \exp \left( -\frac{\gamma z_j \Phi \Phi_j \varepsilon_j}{L} \right) dx}
\]  

(A7)

The PNP theory is the simplest mean field theory because it assumes that the mean potential and concentrations of Poisson’s equation are identical to the mean potential and concentrations in the NP equations. The system of equations (A1) and (A6) or (A7) are self-consistent because they can be (and must be) solved simultaneously. The potential depends on the concentrations \(C_{j,x}\) through the Poisson equation (A11) but the concentrations also depend on the potential through the integrated Nernst-Planck equations (A7). Indeed, the concentrations depend exponentially on the potential profile \(\Phi \Phi_j \varepsilon_j\). Once the potential profile is set, the distribution of concentration is determined and cannot change. In that sense, assuming a single profile \(\Phi \Phi_j \varepsilon_j\) is equivalent to locking a channel into a specific occupancy state, from which it cannot move as long as the potential profile itself does not change.

A different integration of the Nernst-Planck equations shows that flux (and the observable, the electric current \(I\)) also depends exponentially on the potential profile \(\Phi \Phi_j \varepsilon_j\).

\[
I = \frac{D_j C_j(L) \exp(-\gamma z_j V_{applied})}{\int_0^L \exp \left( -\frac{\gamma z_j \Phi \Phi_j \varepsilon_j}{L} \right) dx} - \frac{D_j C_j(R) \exp(-\gamma z_j V_{applied})}{\int_0^L \exp \left( -\frac{\gamma z_j \Phi \Phi_j \varepsilon_j}{L} \right) dx}
\]

(A8)

While the numerators of these terms can be written as functions of (just) the electrochemical potential, the denominators cannot (Chen & Eisenberg, 1993b; Eisenberg, 1995). The denominator and flux itself depends exponentially on \(\Phi \Phi_j \varepsilon_j\), the electrical potential profile itself (not the electrochemical potential), which in turn depends on all the variables and parameters of the system through Poisson’s equation (A1). Once the potential profile is set, the flux is determined. In that sense, assuming a potential profile is equivalent to locking a channel into a specific conducting state, from which it cannot move as long as the potential profile does not change.

Each term of equation (A8) describes the (so-called) unidirectional flux measured by tracers (usually radioactive isotopes) moving into a medium of (nearly) zero tracer concentration (see precise definition of unidirectional flux in Chen and Eisenberg (1993b)), and reference to the extensive literature in, for example, Jacquez (1985). The fluxes can also be written as a chemical reaction (without approximation, for any potential barrier provided concentration boundary conditions are in force)

\[
L \xrightarrow{\text{\(k\)}} R
\]

(A9)

in which the rate constants are the conditional probabilities of the underlying diffusion process described by a full Langevin equation and boundary conditions (Eisenberg, Klosek & Schuss, 1995, Eq. 4) that can be solved if dominates, \(\varepsilon\).

\[
k_j = \text{Prob}(R \rightarrow L) = \frac{D_j \exp(-\gamma z_j V_{applied})}{\int_0^L \exp \left( -\gamma z_j \Phi \Phi_j \varepsilon_j \right) dx}
\]

\[
k_j = \text{Prob}(L \rightarrow R) = \frac{D_j \exp(-\gamma z_j V_{applied})}{\int_0^L \exp \left( -\gamma z_j \Phi \Phi_j \varepsilon_j \right) dx}
\]

(A10)

These expressions are not as obvious as they seem; until this work was done, it was not clear whether a rate constant formulation like Eqs. (A9) and (A10) could be derived for potential barriers of arbitrary shape. The conditional probabilities are in fact specified as the solution of a rather involved partial differential equation and boundary conditions (the full Fokker-Planck equation); showing that took us some seven years (Cooper, Gates & Eisenberg, 1988a; Barcilon et al., 1995; Eisenberg, 1996), only at the end of which was the simple expression (A10) derived.\(^{58}\)

**Numerical Solution**

The design and choice of the numerical procedure for solving PNP is critical for the success of the calculations (Fatemi, Jerome & Osher, 1991; Jerome, 1995). The procedure was developed independently by Chen, e.g., Chen & Eisenberg, 1993a. Severe numerical difficulties arise with other methods of solving the system of Eqs. (A1) and (A6) and boundary conditions (A2) + (A5) (e.g., successive approximation to integral equations. Chen, Barcilon & Eisenberg, 1991; efficient differential equation solvers; see electrochemistry literature, chiefly from Buck and Mafté’s group. op. cit.).

The coupled nonlinear system of equations (A1) and (A6) is solved numerically by substituting the integrated Nernst-Planck equation (A7), which is precisely equivalent to the differential Nernst-Planck equation (A6), into a discretized version of the Poisson equation (A1). We start with a reasonable initial guess, say a discretized version of the constant field potential

\[
\Phi(x; \text{initial guess}) = \left( 1 - \frac{x}{d} \right) \left[ \frac{\ln f}{e} V_{applied} + \varphi_0(0) - \varphi_0(d) \right] + \varphi_0(d)
\]

(A11)

That initial guess of the potential profile is substituted into the right hand side of (a discretized version of) the integrated Nernst-Planck equation (A7) to determine the congruent initial guess of concentration \(C_{j,x; \text{initial guess}}\) and that guess is substituted into the right hand side of Poisson’s equation, which is then solved (it is linear!). The resulting estimate of the potential \(\Phi(x; \text{first iterate})\) is substituted into the integrated Nernst-Planck equation (A7) and so determines a first-iterate of the concentration profiles \(C_{j,x; \text{first iterate}}\). These two first-iterates \(\Phi(x; \text{first iterate})\) and \(C_{j,x; \text{first iterate}}\) are substituted into the right hand side of Poisson’s equation, which is then solved, now to determine the second-iterate \(\Phi(x; \text{second iterate})\), a better approximation to the potential profile. The second-iterate of potential determines a second-iterate of concentration by Eq. (A7); together, the two second-iterates determine the third-iterate of potential, and so on for ten iter-
ations which typically converge in less than one second to better than one part in $10^{-15}$ on our workstation, an IBM RS/6000 Model 550. Each iteration involves a solution of the (discretized version) of Poisson’s equation at 4,000 points of a uniform spatial mesh and yields a potential profile $\Phi(x)$, and concentration profiles $C_j(x)$ for one applied potential $V_{\text{app}}$ and one set of concentrations $C_j(L)$ and $C_j(R)$. From these profiles, the flux $J_j$ and current $I$ at that potential $V_{\text{app}}$ and, for those concentrations, are calculated by Eq. (A8). Repeating the calculation at different potentials produces an IV curve in some 100 seconds of computer time if $I(V_{\text{app}}; C_j(L), C_j(R))$ is determined at 100 values of $V_{\text{app}}$. Of course, a different IV curve is computed for each set of concentrations $C_j(L), C_j(R)$.

The program that executes this numerical procedure is written in FORTRAN 77 and is available to anyone who requests it. The program has compiled and run easily on a number of systems and is a useful, if not necessary, aid to understanding the behavior of the PNP equations.