Basic principles of membrane transport

STANLEY G. SCHULTZ

Basic principles of membrane transport

STANLEY G. SCHULTZ

Professor and Chairman, Department of Physiology
University of Texas Medical School
Houston, Texas

CAMBRIDGE UNIVERSITY PRESS
CAMBRIDGE
LONDON NEW YORK NEW ROCHELLE
MELBOURNE SYDNEY

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
32 East 57th Street, New York, NY 10022, USA
296 Beaconsfield Parade, Middle Park, Melbourne 3206, Australia

First published 1980

Printed in the United States of America
Typeset by Automated Composition Service, Inc., Lancaster, Pennsylvania
Printed and bound by Vail-Ballou Press, Inc., Binghamton, New York

Library of Congress Cataloging in Publication Data

Schultz, Stanley G.

Basic principles of membrane transport.

(IUPAB biophysics series; 2)
1. Biological transport. I. Title. II. Series: International Union for Pure and Applied Biophysics. IUPAB biophysics series; 2.
QH509.S38 574.8'75 79-54015
ISBN 0 521 22992 8 hard covers
ISBN 0 521 29762 1 paperback

FOREWORD

The origins of this series were a number of discussions in the Education Committee and in the Council of the International Union of Pure and Applied Biophysics (IUPAB). The subject of the discussions was the writing of a textbook in biophysics; the driving force behind the talks was Professor Aharon Katchalsky, first while he was president of the Union, and later as the honorary vice-president.

As discussions progressed, the concept of a unified text was gradually replaced by that of a series of short inexpensive volumes, each devoted to a single topic. It was felt that this format would be more flexible and more suitable in light of the rapid advances in many areas of biophysics at present. Instructors can use the volumes in various combinations according to the needs of their courses; new volumes can be issued as new fields become important and as current texts become obsolete.

The International Union of Pure and Applied Biophysics was motivated to participate in the publication of such a series for two reasons. First, the Union is in a position to give advice on the need for texts in various areas. Second, and even more important, it can help in the search for authors who have both the specific scientific background and the breadth of vision needed to organize the knowledge in their fields in a useful and lasting way.

The texts are designed for students in the last years of the standard university curriculum and for Ph.D. and M.D. candidates taking advanced courses. They should also provide a suitable introduction for someone about to begin research in a particular field of biophysics. The Union is pleased to collaborate with the Cambridge University Press in making these texts available to students and scientists throughout the world.

Franklin Hutchinson, Yale University
Watson Fuller, University of Keele
Lorin J. Mullins, University of Maryland
Editors

PREFACE

My initial reaction upon being invited by the Editorial Committee of the International Union of Pure and Applied Biophysics to write a monograph on principles of membrane transport was a mixture of flattery and terror. I was naturally flattered by the invitation but terrified by the seeming enormity of the task and the effect that acquiescence would have on my already overburdened schedule. When this initial reaction gave way to calm objectivity, I asked myself: With so many books appearing annually on the subject of membrane transport, is there a justifiable need for yet another?

My answer to this question, which was confirmed by many of my colleagues, was that there still is a need for the type of "primer" I would have desperately longed for in 1959 when, after completing a Residency in Internal Medicine, I joined the Biophysical Laboratories of the Harvard Medical School as an initiate in the field of membrane transport; my contacts with undergraduates, graduate students, postdoctoral fellows, and junior scientists during the past two decades have reinforced this feeling.

In short, this monograph is intended for the "initiate." Its goal is to introduce some of the basic principles that govern solute and water transport across membranes and to emphasize the foundations and reasoning that underlie these principles in a relatively brief and readable form. It is not intended to supplant the more advanced and comprehensive treatments that can be found in a number of books, chapters, and review articles, but rather to provide the readers with an appreciation of basic concepts and approaches that will ease their way into the more advanced literature. In order to achieve this goal, I have attempted to develop this monograph from "first principles" assuming only that the reader has had some previous exposure to elementary differential and integral calculus, chemistry, and physics. All final expressions are derived, inasmuch as possible, step by step, and particular care has been taken to see to it that a consistent set of conventions is carefully defined and adhered to throughout. There are a number of basic points that are repeated throughout this monograph; this may irritate the more advanced reader but, at the same time, could be valuable to the beginner. In any event, this introductory text is designed to teach and I am convinced that repetition is an essential ingredient of pedagogy.

If I have succeeded in achieving my intended goal, a large measure of the success is owed to my teachers, colleagues, and students.

I am particularly indebted to:

Professor Arthur K. Solomon, who opened the door to my scientific career and who permitted (urged) me to teach the first offering of a course entitled "Bioenergetics and Membrane Transport" (Biology 119) at Harvard College in 1967. This was a memorable experience which forced me to come to grips with theoretical principles from the vantage point of a teacher rather than a user. Much of the organization and many of the developments of this text are based on my lecture notes for that course.

The late Professor Peter F. Curran, a long-time combination teacher-collaborator-friend, who was exceptionally capable of employing theory to provide a logical, coherent, and concise "picture" of experimental data.

And, the late Professor Aharon Katchalsky, a profound thinker, a gifted educator, and a constant source of inspiration; one cannot lavish too much praise upon his personal and intellectual attributes.

The tragic and untimely deaths of Peter Curran and Aharon Katchalsky, within a brief span of time, came as terrible blows. I am privileged to have shared their friendship and collegiality, and owe them both a great deal.

In addition, I am grateful to Drs. Alvin Essig, Raymond Frizzell, and Stephen Thompson for their careful reading of this manuscript and their corrections, cogent suggestions, and critical comments.

I wish to thank my secretary, Susan Frizzel, who nurtured this effort from first draft to final product and am eternally indebted to her dog, Fagin, for not having completely devoured Chapter 4 during an unguarded moment. (I reject my son's suggestion that he found it "too dry"!)

Finally, I owe a special brand of gratitude to my wife, Harriet, and my children, Jeffery and Kenneth, who have always provided me with the warmth, encouragement, and peace of mind to enable me to "do my thing." My debt to them and my parents can never be repaid.

S.G.S.

CONTENTS

1 Some basic principles of thermodynamics: The relations between flows and forces

The formal description of transport across membranes rests firmly on the unassailable citadel of thermodynamics, that branch of science concerned with the changes that accompany the transition of a system (a defined portion of the universe selected for study) from one set of conditions (state) to another.[1] In this chapter we will develop the basic principles that deal with the transitions of a system from one state (initial) to another (final), which result from the transfer of matter within the system.[2]

1.1. Properties of state

Any macroscopic system at equilibrium can be described completely by an *equation of state*, which relates a set of parameters or properties that are functions only of the current state of the system and are independent of its past history. It follows that any change in a *property of state* is uniquely determined by the initial and final conditions of the system and is independent of the path (or mechanism) of transition between these two states.

A *property of state* can be formally defined as follows. If Z is a function of other properties of state, for example, $Z = f(U, V, W, \ldots)$ and if

$$dZ = L \cdot dU + M \cdot dV + N \cdot dW + \cdots$$

where $L = (\partial Z/\partial U)_{V,W,\ldots}$, $M = (\partial Z/\partial V)_{U,W,\ldots}$, $N = (\partial Z/\partial W)_{U,V,\ldots}$, and so on, then dZ is an *exact* (*or total*) *differential* of $f(U, V, W, \ldots)$ and Z is a property of state.[3] The reason, of course, is that if these relations hold, then $L \cdot dU$, $M \cdot dV$, $N \cdot dW$, and so on, represent the *individual contributions* of the changes in $U, V, W, \ldots$ to the *total change* in Z and, since addition is a commutative operation, the total change in Z does not depend upon the order or sequence of the changes in $U, V, W, \ldots$. For example, referring to Figure 1.1, if a system undergoes a change from state A to state C as a result of changes in properties U and V, the change in Z will be the same if the change in U preceded the change in V (pathway 1) or if both U and V change simultaneously (pathway 2), that is,

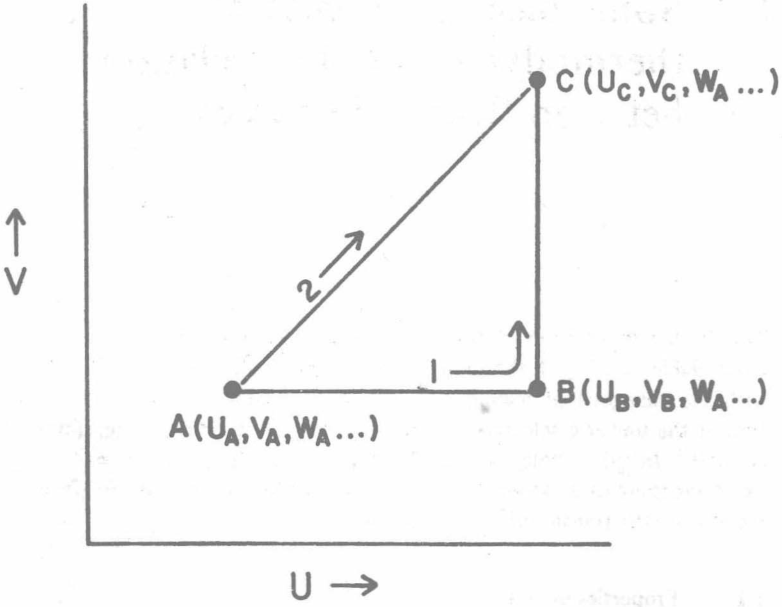

Figure 1.1. $U, V, W, \ldots$ are properties of state. If a system undergoes a transition from an initial state A to a final state C as a result of changes in U and V, the properties of the final state are independent of the path.

$$\begin{aligned}\Delta Z = Z_C - Z_A &= f(U_C, V_C, W_A, \ldots) - f(U_A, V_A, W_A, \ldots)\\ &= [(Z_B - Z_A) + (Z_C - Z_B)]\end{aligned}$$

Finally, the fact that dZ is an exact differential means that there will be no change in Z if the system is displaced from state A and then restored to state A regardless of the paths (mechanisms and durations) of displacement and restoration (e.g., referring to Figure 1.1, A to B and back, A to C and back, A to B to C and back, A to C to B and back); that is, the closed integral, $\oint dZ = 0$ (Green's theorem).

The state properties of a homogeneous system can be divided into two categories, namely, *intensive properties* and *extensive properties*. Intensive properties are those characteristic of the total system, as well as any arbitrarily selected part of the system, and they are nonadditive; for example, pressure (P), temperature (T), concentration (c), and electrical potential (ψ); these properties reflect the "intensity" of the system. The extensive properties of a homogeneous system are measures of the size or content of the system and are additive; for example, volume (V), number of moles (n). Obviously, the total volume of a system is the sum of the volumes of all sub-

Table 1.1. *Conjugate properties of state*

Intensive property	Extensive property
Pressure (P)	Volume (V)
Electrical potential (ψ)	Charge (e)
Temperature (T)	Entropy (S)
Chemical potential (μ)	Moles (n)

divisions (parts) of the system whereas the temperature of these subdivisions is the same. Further, it is obvious that the ratio of any two extensive properties is an intensive property (e.g., $n/V = c$), and it can be shown that any intensive property can be expressed as a function of all other intensive properties of a system (e.g., the perfect gas law $PV = nRT$ can be written $P = cRT$).

Every intensive property is paired with a *conjugate* extensive property and a difference in an intensive property is *a* driving force for the flow or displacement of its conjugate extensive property. The conjugate properties of state with which we will be concerned are listed in Table 1.1. Clearly, a difference in pressure (ΔP) is a driving force for the displacement of volume (V); a difference in electrical potential ($\Delta\psi$) is a driving force for the flow of charge (e); as we will see, a difference in temperature (ΔT) is a driving force for the displacement of entropy (S); and for the moment let us define a driving force for the flow of uncharged matter (n) as a difference in chemical potential ($\Delta\mu$). An explicit expression for the chemical potential will be derived below. Further, for reasons that will become evident, we refer to a difference in an intensive property as the *conjugate driving force* for the flow of its paired (conjugate) extensive property. Finally, it should be noted that the product of an intensive property and its conjugate extensive property has units of energy or work. Thus, $P\,dV$ is the work that must be performed to displace an amount of volume (dV) against a pressure (P); $\psi\,de$ is the work that must be performed to bring an amount of charge (de) from a region of zero electrical potential (ground state) to a region where the electrical potential is ψ; and as we will see, $\mu_i\,dn_i$ is the work that must be performed to transfer a number of moles of an uncharged substance (dn_i) from a region where the chemical potential of i is zero to one where it is μ_i.

1.2. The first and second laws of thermodynamics

The entire edifice of thermodynamics is founded on two laws or postulates that simply summarize a vast amount of human experience. The *first law*, often referred to as the law of conservation of energy, states that the total internal energy of a system (E) can only change as a result of the gain or loss of heat (Q) and/or performance of work (W) on or by the system;

that is,

$$dE = dQ - dW \tag{1.1}$$

where dQ is positive when heat is gained by the system and dW is positive when work is performed by the system on its surroundings. E is an extensive property of state determined by the total potential and kinetic energy of the components within the system, and dE is an exact differential of an equation of state. However, Q and W are not properties of state inasmuch as the amount of heat gained and work performed by a system when it changes from one state to another depends upon the nature (pathway, mechanism) of the transition.

For a system that is capable of exchanging matter with its surroundings, equation (1.1) can be written more explicitly as follows:

$$dE = dQ - P\,dV + \psi\,de + \sum_i \mu_i\,dn_i \tag{1.2}$$

where $P\,dV$ is positive when the system performs work by expanding (dV) against an external pressure P; $\psi\,de$ is positive when an amount of charge de is transferred into a system whose electrical potential is ψ; and $\mu_i\,dn_i$ is positive when an amount (moles) of matter dn_i is transferred into a system where the chemical potential of i is μ_i (see note 9).

The first law is a "bookkeeper's delight"; it states that a change in the internal energy of a system must be the sum of the exchange of heat with its surroundings and the work performed on or by its surroundings. Since the heat *gained* or *lost* by the system must be equal to the heat *lost* or *gained* by the surroundings, and since the energy lost by a system as a result of the performance of work on its surroundings must be equal to the energy gained by the surroundings, the internal energy of the universe is conserved; E is simply transferred from one system to another but the total remains constant. However, the first law provides no way of distinguishing between spontaneous or natural processes and processes that can never take place spontaneously; according to this law, all processes are equally possible providing they obey equation (1.1).

The *second law* of thermodynamics is the culmination of many efforts to describe the fact that all natural or spontaneous processes take place only in the direction toward equilibrium and come to a halt when this time-independent state is reached! Work must be performed on a system to displace it from equilibrium; such a displacement cannot occur spontaneously and, for this reason, spontaneous processes are referred to as "irreversible." This monumental achievement was accomplished by introducing the concepts of reversible processes, namely, processes that take place infinitely slowly so that the transition from one state to another may be treated as if the system passes through an infinite series of equilibrium states. For such an idealized process, it can be shown that, although dQ *is not* an exact differential, the ratio (dQ/T) *is* an exact differential. Thus, a new property of state emerges, the entropy (S), which for a reversible process is defined by the relation

$$dS = dQ/T \tag{1.3}$$

Since S is a property of state, a change in S is dependent only on the initial and final states of the system and is independent of the intervening path. Thus, for a given set of initial and final conditions, dS is the same if the transition between these states is the result of a spontaneous (irreversible) process or an idealized reversible process. However, for a given transition, the heat lost to the surroundings is greater (or the amount of heat taken up by the system from its surroundings is less) when the process is irreversible than when it is reversible.[4] Thus, for an irreversible process

$$dS > dQ/T \tag{1.4}$$

Stated another way, if the same change in state is carried out reversibly and irreversibly

$$dS = (dQ/T)_{\text{rev}} > (dQ/T)_{\text{irrev}}$$

That is, the heat that is actually absorbed in any real process is less than what would have been absorbed had the process been reversible. The difference between $(dQ/T)_{\text{rev}}$ and $(dQ/T)_{\text{irrev}}$ was referred to by Clausius as "the uncompensated heat."

Equations (1.3) and (1.4) for reversible and irreversible processes, respectively, comprise a formal statement of the second law of thermodynamics.

The direction of a spontaneous or irreversible process must satisfy the inequality given in equation (1.4). Since S is an extensive property of state, dS can be divided into two parts, that is,

$$dS = d_eS + d_iS$$

where d_eS represents the change in entropy of a system caused by exchange of heat with the surroundings and d_iS represents the internal production or creation of entropy (the "uncompensated heat"). For an irreversible process $d_eS + d_iS > dQ/T$, where $d_eS = dQ/T$, so that $d_iS > 0$; for a reversible process $d_eS + d_iS = dQ/T$, so that $d_iS = 0$. Thus, an irreversible process can be defined as one that leads to an internal production of entropy that cannot be accounted for by an exchange of heat with the surroundings. Further, it should be noted that for an irreversible process, d_eS and dS can be positive or negative depending upon whether heat is gained or lost by the system, but d_iS is always positive.

Finally, since E is a property of state, dE is independent of path and for a given change in state is the same whether the transition is the result of a reversible or an irreversible process. Thus, we can substitute $T\,dS$ for dQ in equation (1-2) and obtain

$$dE = T\,dS - P\,dV + \psi\,de + \sum \mu_i\,dn_i \tag{1.5}$$

From this equation, it is clear that S is the extensive conjugate of T in the displacement or flow of thermal energy.[5]

Equation (1.5) is often referred to as the "Gibbs equation," and in its most general form it includes changes in all possible extensive properties and equates dE to the sum of the products of intensive properties and changes in their conjugate extensive properties, that is,

$$dE = \sum_i [(\text{Intensive property})_i \times (\text{Change in conjugate extensive property})_i]$$

1.3. The Gibbs free energy and the concept of "useful work"

Although the entire fabric of classical thermodynamics can be woven out of the first and second laws without introducing any additional properties of state, equation (1.5) does not provide the most convenient or illuminating framework for the thermodynamic description of processes that take place under conditions of constant temperature and pressure. To remedy this shortcoming, Gibbs introduced the free-energy function G, which is defined as

$$G = E + PV - TS \tag{1.6}$$

Since E, P, V, T, and S are properties of state, G must also be a property of state, so that dG is determined entirely by the initial and final conditions of a system and is independent of the intervening path. From equation (1.6)

$$dG = dE + P\,dV + V\,dP - T\,dS - S\,dT \tag{1.7}$$

and when T and P are constant

$$(dG)_{T,P} = dE + P\,dV - T\,dS \tag{1.8}$$

We now (arbitrarily) divide the total work done by (or on) a system into "pressure × volume work" and "useful work" so that

$$dW = P\,dV + dW' \tag{1.9}$$

where dW' represents the useful work and, as before, is positive when the system performs work on its surroundings and is negative when work is performed on the system.[6]

Substituting equation (1.9) into equation (1.1), we obtain

$$dE = dQ - P\,dV - dW' \tag{1.10}$$

and substituting equation (1.10) into equation (1.8) we obtain

$$(dG)_{T,P} = dQ - T\,dS - dW' \tag{1.11}$$

Since $dS = d_eS + d_iS = (dQ/T) + d_iS$,

$$T\,dS = dQ + T\,d_iS$$

so that from equation (1.11) we obtain

$$-(dG)_{T,P} = T\,d_iS + dW' \qquad (1.12)$$

Thus:

(a) If a system at constant T and P performs *useful work* on its surroundings by means of a reversible process, $d_iS = 0$ and the decrease in the Gibbs free energy of the system ($-dG$) is equal to the useful work performed. If the process is irreversible, $d_iS > 0$ and the decrease in G is *greater* than the useful work performed.

(b) If a spontaneous process occurs in a closed system (one that cannot exchange matter with its environment), $dW' = 0$ and the decrease in free energy of the system is reflected only by an increase in internal entropy; that is, free energy is completely dissipated.

These considerations provide some insight into the meanings of internal energy, free energy, and entropy. Consider for the moment an irreversible process that takes place in a closed system (e.g., the diffusion of a solute from a region of higher concentration to one of lower concentration). Clearly, the internal energy of the system will not be affected since dQ and dW are both zero. However, the free energy of the system will decrease and reach zero when equilibrium is achieved. Thus, the potential ability of the system to perform work is dissipated at the expense of the production of entropy; for this reason free energy functions are frequently referred to as "work functions." The same holds for a system that can perform work on its surroundings. For any real process the decrease in free energy is greater than the amount of useful or recoverable work performed. Thus, the free energy of the universe has decreased but the internal energy is conserved.

Why? At the outset it should be stressed that thermodynamics is not based on any particular model of matter, and the functions E and S were introduced as properties of state that simply permit a consistent description of the energetics and direction of natural events. We cannot search for a more profound interpretation of these functions in classical thermodynamics. However, let us depart from this secure axiomatic approach and consider a closed system subdivided into two compartments by a porous barrier with one compartment containing a monoatomic gas and the other initially void. Clearly, the gas will diffuse (expand) into the initially empty compartment and equilibrium will be achieved when the distribution of atoms is random (i.e., equal concentrations in both compartments). Now, the internal energy of this system is simply the sum of the kinetic, translational energies of the individual atoms and is given by $E = n \cdot m \cdot \overline{v}^2/2$ where n is the number of atoms, m is the atomic mass and $\overline{v}^2$ is the mean square velocity of each atom that is only dependent on T. It should be noted that although $\overline{v}$ is a vector, $\overline{v}^2$ is a scalar, so that while E embodies the concept of motion, it is divorced from the notion of direction. E would be the same if at any instant every atom were moving in the same direction as when the movements are directed randomly. In contrast, the entropy function is a measure of the randomness of the system. An increase in entropy or an increase in randomness may be defined as an increase in the number of possible configurations or complexions of a system, all of which are equally likely and accessible; the state of equilibrium represents the most random or most probable configuration that the system can achieve. Thus, the second law states that the direction of all spontaneous processes is from a less random (less probable) to a more random (more probable) state. Clearly, with an increase in the randomness of a system, the ability of the system to convert its internal energy to a directed effort, or useful work, decreases.[7]

At the risk of belaboring this point, let us close this discussion with a trivial example. When a heavy weight is dropped onto a metal plate in a vacuum, the kinetic energy of the falling weight is completely converted into thermal energy, which in turn is reflected by an increase in the velocity of random motion of particles (entropy). If the motion were not random, but instead if all of the particles could be harnessed to simultaneously move upward against the weight, the weight would rise and a spontaneous process will have been reversed without outside intervention. The second law infers that the *probability* of this occurrence is so low that this event may be considered impossible.[8]

1.4 The electrochemical potential

We now employ the Gibbs free energy function to derive an expression for the electrochemical potential of a substance i. As we will see, the electrochemical potential is an intensive property of a system, and a difference in electrochemical potential is *the* conjugate driving force for the diffusional flow of matter. This property will play a central role in all of our subsequent considerations so that the importance of understanding its meaning cannot be overstated.

Let us start by considering the effect of reversibly transferring matter into a system on G. Inasmuch as

$$G = E + PV - TS \tag{1.6}$$

$$dE = dQ - P\,dV - dW' \tag{1.10}$$

and for a reversible process

$$dQ = T\,dS$$

it is a simple matter to show that

$$dG = -S\,dT + V\,dP - dW' \tag{1.13}$$