



Introduction to Biophysical
Methods for Protein and
Nucleic Acid Research

EDITED BY

Jay A. Glasel and Murray P. Deutscher



Introduction to
Biophysical Methods
for Protein and
Nucleic Acid Research

Edited by

Jay A. Glasel

Murray P. Deutscher

Department of Biochemistry
University of Connecticut Health Center
Farmington, Connecticut



Academic Press

San Diego New York Boston London Sydney Tokyo Toronto

Cover photo: The β subunit of the *E. coli* replicase, DNA polymerase III holoenzyme, encircles DNA acting as a sliding clamp to tether the rest of the replicase machinery to the chromosome. The two subunits of the holoenzyme are shown in red and yellow. The commercial molecular modeling program "Quanta" (Molecular Simulations, Inc., Burlington, MA) operating on a Silicon Graphics, Inc., Indigo computer was used to produce the Postscript output. For other views of the same structure see Kong, X.P., Onrust, R., O'Donnell, M., and Kuriyan, J. (1992), *Cell* 69 425–437. Courtesy of Drs. X. P. Kong and J. Kuriyan, Laboratory of Molecular Biophysics, The Rockefeller University, New York, N.Y., and Drs. R. Onrust and M. O'Donnell, Microbiology Department, Cornell University Medical College, New York, N.Y.

This book is printed on acid-free paper. ∞

Copyright © 1995 by ACADEMIC PRESS, INC.

All Rights Reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Academic Press, Inc.

A Division of Harcourt Brace & Company
525 B Street, Suite 1900, San Diego, California 92101-4495

United Kingdom Edition published by
Academic Press Limited
24-28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in-Publication Data

Introduction to biophysical methods for protein and nucleic acid research /
edited by Jay A. Glasel, Murray P. Deutscher.

p. cm.

Includes bibliographical references and index.

ISBN 0-12-286230-9

1. Proteins—Research—Methodology. 2. Nucleic acids—Research—
-Methodology. 3. Amino acid sequence. 4. Nucleotide sequence.

I. Glasel, Jay A. II. Deutscher, Murray P.

QP551.I63 1995

574.19'245—dc20

95-8162

CIP

PRINTED IN THE UNITED STATES OF AMERICA

95 96 97 98 99 00 QW 9 8 7 6 5 4 3 2 1

**Introduction to Biophysical Methods
for Protein and Nucleic Acid Research**

Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

- Norma M. Allewell** (381) Department of Biochemistry, College of Biological Sciences, University of Minnesota, St. Paul, Minnesota 55108
- B. W. Bangerter** (317) Department of Chemistry, Yale University, New Haven, Connecticut 06520
- Michael B. Bolger** (433) University of Southern California School of Pharmacy, Los Angeles, California 90033
- Richard M. Caprioli** (147) The Analytical Chemistry Center and Department of Biochemistry and Molecular Biology, The University of Texas Medical School, Houston, Texas 77030
- David E. Garfin** (53) Life Science Group, Bio-Rad Laboratories, Inc., Hercules, California 94547
- Jay A. Glasel** (1) Department of Biochemistry, University of Connecticut Health Center, Farmington, Connecticut 06032
- Arthur R. Hand** (205) Department of Pediatric Dentistry, School of Dental Medicine, and Central Electron Microscope Facility, University of Connecticut Health Center, Farmington, Connecticut 06030
- Takashi Miura** (261) Pharmaceutical Institute, Tohoku University, Aobayama, Sendai 980, Japan
- Todd M. Schuster** (111) Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06268
- Walter F. Stafford** (111) Boston Biomedical Research Institute, Boston, Massachusetts 02114
- Marc J.-F. Suter** (147) Chemistry Department, Swiss Federal Institute for Environmental Science and Technology (EAWAG), 8600 Duebendorf, Switzerland
- George J. Thomas, Jr.** (261) School of Biological Sciences, University of Missouri—Kansas City, Kansas City, Missouri 64110
- Jaishree Trikha** (381) Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Preface

Driven by developments in precision instrumentation, in electronics miniaturization, and in computer hardware and software, and by the growing interest in understanding structures of biomolecules, biophysical methods have entered the working lives of molecular biologists and biochemists in a forceful way. To do state-of-the-art research in many areas of biological science, a student or senior scientist must be aware of what instruments are available, what information can be obtained, what alternatives exist, etc. As an example of just how important such methods can be, most of us routinely use gel filtration chromatography to obtain molecular masses of proteins and nucleic acids with accuracies approaching $\pm 10\%$ (neglecting effects of shape). However, mass spectrometric methods have now been developed that can yield molecular masses of macromolecules to $\pm 0.001\%$, and in less time. Yet, many graduate students and researchers in the biological sciences lack the mathematical and physical backgrounds necessary to use biophysical procedures fruitfully, and, often, they are even unaware of biophysical tools that they might effectively employ in their research. While many advanced monographs and textbooks that deal with biophysical methods are available, they generally do not serve the needs of those requiring an introduction to the field.

This book is meant to occupy niches both as a textbook and as an initial reference source. Not only will it serve the needs of advanced undergraduates and graduate students learning biophysical techniques for the first time, but it also has been designed to provide experienced investigators with enough information about biophysical procedures to enable them to carry out a method on their own or to collaborate with an expert in the field. In neither case do we assume a working knowledge of the mathematics of physical chemistry. The subject areas have been chosen to cover a wide range of the methods most frequently needed to answer current biological problems. Each chapter presents a description of the physical basis of the method, the type of information that may be obtained with the method, how data should be analyzed and interpreted, and, where appropriate, practical tips about equipment and procedures. Mathematics is included at a level to make the methods understandable and useful without being necessarily complete and rigorous. Thus, we hope to provide the basic understanding that will enable a nonphysically oriented re-

searcher to begin to speak in a common language with an experienced practitioner. It is our goal to open the field of biophysical methodology to those who might otherwise not appreciate the importance and usefulness of this area for their own biological training or research problems.

Biophysical methods are now so numerous that this book makes no pretense of being all-inclusive. We have attempted to include those methods that are in wide use and whose understanding is fundamental to learning about new developments. Other methods are not discussed although they are widely used and were once at the forefront of experimental biophysics. For example, molecular biology could not have advanced to its present stage without the widespread use of radioactive tracers. However, molecular biologists rarely consider the detailed physics behind their use. Furthermore, we are entering an era in which the use of newer nonradioactive methods is rapidly replacing beta and gamma counters. Thus, we have not included a chapter on biophysical applications of radioactivity. On the other hand, some techniques as old as the modern era of physical chemistry are currently undergoing revivals because of advances in instrumentation, and for such methods we have included discussion.

In contrast to most biophysical methods, there is one that does not even involve direct measurements on molecules. That tool is computer molecular modeling and graphics. It was not long ago that in some universities computer-generated images of objects were under the administration of art departments. While that era may not be distant in terms of years, in technical terms it is as far away as the middle ages. Tremendous desktop computing power is now available in thousands of laboratories, and computer molecular modeling has emerged as a valuable tool for both research and teaching. The sophistication of modern software coupled with advances in every aspect of computer hardware has brought this about. In addition, many laboratories have physical links between their desktop and remote computers that can serve and receive data and programs. That is, most of us are connected to "the network." Because of the importance and increasing use of the methodologies, a chapter describing computer applications has been included.

Inasmuch as a scientific language barrier often exists that prevents many students and working scientists from appreciating what biophysical specialists are talking about, each chapter includes a glossary of terms and concepts. The glossaries contain definitions and sometimes extensive discussion. They are meant for readers who are unfamiliar with the concepts of physical biochemistry. Depending on the subject, more advanced readers may want to read through a chapter with little, or only occasional, reference to the Glossary for that chapter. Glossary items are indicated in boldface type the first time they appear in each chapter.

The first, introductory, chapter summarizes the basic properties of macromolecules. It discusses the mass, shape, electrical, and magnetic properties of proteins, nucleic acids, and their components. In particular, this chapter provides the reader with the scientific vocabulary used throughout the book. In addition, it indicates to the reader what aspects of macroscopic behavior are complex and what aspects are simple.

The next two chapters concentrate on descriptions of modern versions of venerable techniques used to determine molecular masses, shapes, and interac-

tions. Thus, the second chapter, "Electrophoretic Methods," presents a discussion of the modern manifestation of electrophoresis. While almost all biochemists now do some type of electrophoretic measurement in their laboratories, they may not be aware of the breadth of electrophoretic methods and available detection methods. In this chapter, the poorly understood theory of electrophoresis is overshadowed by the empirical methods developed to do electrophoretic experiments. Consequently, this chapter is much more of a "how-to" discussion than any other in the book.

The third chapter deals with analytical ultracentrifugation and viscosimetry. Other than in mass spectrometry, nowhere are the new technical developments more evident than in ultracentrifugation. Those of us familiar with the Beckman "Model E" ultracentrifuge, and its bulk and instrumental complexity, marvel at developments that have shrunk the size of the ultracentrifuge to that of a 3' × 3' × 3' preparative ultracentrifuge. Moreover, the new instrument is vastly more "user-friendly," and so is the associated computer analysis of the data. On the other hand, the basic explanations of the behavior of macromolecules in an analytical ultracentrifuge have not changed over the years. In Chapter 3, the description of ultracentrifugation is brought out in a simple way, and is followed by current applications of the method. Viscosimetry is also included because it is an often-neglected, simple technique that can give quick results on certain overall macromolecular properties (e.g., deviations from spherical shape).

Chapter 4 describes the theory and instrumentation underlying the exciting developments that have taken place in mass spectrometry within the past ten years. While the theory of mass spectrometry is quite simple, the basis for the production of ionized macromolecules in a vapor phase is certainly neither simple nor completely understood. Consequently, ionization techniques and their respective applications to biophysical measurements are described rather than theoretically analyzed. This chapter presents an extensive discussion of applications of modern mass spectrometry that are of interest to biochemists and molecular biologists. These include protein molecular mass measurements, determination of posttranslational modifications, and the rapidly developing strategies for peptide sequencing.

Chapter 5 describes the application of electron microscopic methods to the determination of macromolecular shapes. Electron microscopy has historically been of great use in visualizing overall macromolecular shapes, particularly those of nucleic acids. More recently, electron micrographic images of protein assemblies have been used, via Fourier transform techniques, to obtain macromolecular structures. The fact that this mathematical technique is brought up in two other chapters (Chapters 7 and 8) underscores the necessity for beginners to learn something about fundamental, important, theoretical concepts that can be applied to a variety of measurements.

Chapter 6 is the first of two chapters devoted to spectroscopic methods. The basic aspects underlying all spectroscopy are dealt with in Chapter 1. However, Chapter 6 focuses on the application of the theory to electronic and vibrational spectroscopy. The section on electronic spectroscopy deals with measurements that are done regularly in many laboratories, for example, concentration determination of proteins and nucleic acids. However, it also deals with the resurgence in a technique—circular dichroism—that fell from favor so completely

that many younger workers may not know about it, but which now can be very important for assessing protein secondary structure. The discussion also includes applications of fluorescence spectroscopy to biophysical problems. The chapter then goes on to discuss vibrational spectroscopic techniques (Raman scattering, infrared absorption) that can provide information on intramolecular structural details in proteins and nucleic acids as well as information on kinetics of isotopic exchange. This section includes a discussion of the newly emerging field of ultraviolet resonance Raman spectroscopy and its application to proteins and nucleic acids.

Chapter 7, the second of the spectroscopic chapters, discusses applications of nuclear magnetic resonance (NMR) techniques to biochemistry and molecular biology. This chapter forms a bridge between low and medium resolution techniques for determining molecular structure (as discussed in previous chapters) and high resolution ones such as X-ray and neutron diffraction measurements (discussed in Chapter 8). Thus, low resolution techniques such as electron microscopic methods can yield some structural information at resolutions of several angstroms. A medium resolution technique such as NMR can be used at present to determine *families* of structures of macromolecules in solution with molecular masses less than $\approx 25,000$ daltons. However, X-ray and neutron diffraction can give atomic positions for carbon and higher atomic number atoms to a few tenths of an angstrom in very large, crystallized macromolecules.

NMR theory is presented at a beginning level in Chapter 7 with the object of providing the reader with both scientific concepts and vocabulary. The basic NMR experiment can be understood without a detailed knowledge of quantum mechanics. Thus, this chapter places emphasis on giving the reader a basic outline of what different parameters are derived from NMR data and how they are interpreted in terms of molecular structure. In the current era, multidimensional NMR measurements occupy a central role in macromolecular structure determination. The author of this chapter has tried to present a description of multidimensional NMR in a way that a beginner can get a physical picture of what the experiment involves. He then presents typical examples of successful applications of the technique, with particular attention to the magnitude of a project of structural determination.

Chapter 8 presents the theory of X-ray and neutron diffraction with as little dependence on mathematics as we thought possible for the chapter to be useful. Furthermore, at least two mathematical concepts, the Fourier transform and phase differences in electromagnetic waves, are discussed in other contexts elsewhere in the book. Therefore, the reader should already be comfortable with the concept. The appearance of this mathematical operation in several chapters of the book emphasizes our basic pedagogical philosophy that physical methods should be approached from a unified fundamental background. For example, all forms of spectroscopy have much in common, and X-ray and neutron diffraction phenomena have much in common with their analogous optical effects. As the database of successfully solved macromolecular structures becomes larger, comparative methods of getting starting structures for new X-ray data are becoming more important, and one such method is emphasized in Chapter 8. The chapter also concentrates on explaining aspects of diffraction theory and experiments that most puzzle beginners in the field. For example, the authors have worked hard at explaining the concepts of the reciprocal lattice and the meaning of resolution in a structural determination.

The final chapter of the book concentrates on two separate aspects of using computers, particularly desktop computers, in protein and nucleic acid research. One aspect is obtaining and interpreting structural information. The other is the current state of predicting macromolecular structure. The former is of great interest because of hardware and software advances that have made it possible to access large structural and sequence databases. With such access, it is possible for workers to compare their own structural information with that of previous structures. The author of this chapter has concentrated on how the database information is obtained over the network and how the data so obtained are analyzed.

Theory and practice in prediction of aspects of macromolecular structure are also discussed at a beginning level in this chapter. Many of these methods are being used in ways that may seem insidious to a beginner. For example, as discussed in Chapter 8, after a structure has been "solved" using all available X-ray data, it has become customary to subject the structure to energy minimization on a computer to "anneal" the structure into a more physically realistic one. In some cases, this may move many atomic centers by angstrom distances. The same methods are being used to determine model conformations of, for example, antigen-antibody interactions. In this case both components are annealed using a theoretical treatment and then molecular modeling is used to fit the two structures together.

A special feature of the book is the inclusion of supplementary Macintosh and MS-DOS/Windows disks. These disks contain the molecular modeling programs MAGE and RasMol along with sample data files used to obtain some of the figures in Chapter 9. These programs can be used to build and display new structures or display structural data obtained from databases such as the Protein Data Bank. Therefore, the programs pertain to data obtained by X-ray diffraction, neutron diffraction, and NMR spectroscopy. In addition, several journals now store structural information from published articles in their own databases. These may be obtained via the Internet and displayed using MAGE and RasMol. We have found that contemporary students adapt rapidly to the use of these programs and gain valuable structural insights from the calculations, such as intermolecular distances and angles, available within the programs.

It is hoped that this volume will prove useful to both students and established investigators, and that it will foster greater interaction between the biochemical/molecular biological and the biophysical communities.

We thank the class of 1994, "Biophysical Methods 366, University of Connecticut Health Center," Yingqun Huang, Steve Wiltshire, and Debbie Sardinha for their critiques and suggestions for selected chapters.

*Jay A. Glasel
Murray P. Deutscher*

Contents

Contributors	ix
Preface	xi

Chapter 1 Basic Physical Properties of Proteins and Nucleic Acids

Jay A. Glasel

Glossary	1
I. Introduction	10
II. A History of the Concept That Molecules Can Be Very Large	10
III. Basic Properties of Macromolecules	12
IV. Fourier Transforms	48
V. Summary	51
References	51

Chapter 2 Electrophoretic Methods

David E. Garfin

Glossary	53
Symbols	55
I. Introduction	56
II. Gel Electrophoresis	57
III. Isoelectric Focusing	86
IV. Immunoelectrophoresis	92
V. Blotting	94
VI. Pulsed-Field Gel Electrophoresis	97
VII. Capillary Electrophoresis	100
References	103

Chapter 3 Hydrodynamic Methods*Walter F. Stafford and Todd M. Schuster*

Glossary	111
Symbols	115
I. Introduction	118
II. Analytical Ultracentrifugation	119
III. Sucrose Density Gradient Sedimentation	132
IV. Diffusion	133
V. Sedimentation–Diffusion Equilibrium	134
VI. Viscosity Measurements	140
VII. Resources	143
References	143

Chapter 4 Mass Spectrometry*Richard M. Caprioli and Marc J.-F. Suter*

Glossary	147
I. Introduction	152
II. Instrumentation	153
III. Combined Techniques	172
IV. Applications	174
V. Conclusions	201
References	201

Chapter 5 Electron Microscopy*Arthur R. Hand*

Glossary	205
I. Introduction	208
II. Principles of Electron Optics	209
III. Design and Operation of Electron Microscopes	214
IV. Methods of Sample Preparation	222
V. Applications of Electron Microscopy in Cell and Molecular Biology	244
References	258

Chapter 6 Optical and Vibrational Spectroscopic Methods*Takashi Miura and George J. Thomas, Jr.*

Glossary	261
Symbols	266
I. Introduction	267
II. Electronic Spectroscopy	269
III. Vibrational Spectroscopy	286
Appendix	310
References	313

Chapter 7 Nuclear Magnetic Resonance*B. W. Bangerter*

Glossary	317
I. Introduction	328
II. Basic Principles of Nuclear Magnetic Resonance Spectroscopy	331
III. Experimental Methods for Structure Determination	348
IV. Nuclear Magnetic Resonance of Proteins	359
V. Nuclear Magnetic Resonance of Nucleic Acids	369
VI. Conclusion	376
References	377

Chapter 8 Diffraction Methods*Norma M. Allewell and Jaishree Trikha*

Glossary	381
Symbols	388
I. Introduction	388
II. Crystals	391
III. Collecting Diffraction Data	400
IV. Calculating Electron Density and Patterson Maps	411
V. Obtaining Phases	415
VI. Building Models	422
VII. Refinement	423
VIII. Evaluating the Model	425
IX. Analyzing the Model	426

X. Neutron Diffraction	426
XI. Fiber Diffraction	428
XII. Databases	430
XIII. New Directions	430
References	431

Chapter 9 Computational Techniques in Macromolecular Structural Analysis

Michael B. Bolger

Glossary	433
Acronyms	438
I. Introduction	439
II. Sequence Homology	439
III. Secondary Structure Predictions of Proteins	451
IV. Computers and Graphical Representations	461
V. Calculation of Structural, Thermodynamic, and Catalytic Properties	471
References	487

Index	491
Computer Program Supplement	507
Instructions for Using the Disk Supplements	508

Chapter 1

Basic Physical Properties of Proteins and Nucleic Acids

Jay A. Glasel

Glossary	1	C. Electrical Properties	30
I. Introduction	10	D. Magnetic Properties	38
II. A History of the Concept That Molecules Can Be Very Large	10	E. Kinetic Properties	40
III. Basic Properties of Macromolecules	12	F. Color	43
A. Mass	12	IV. Fourier Transforms	48
B. Shape and Structure	13	V. Summary	51
		References	51

GLOSSARY

Absolute temperature Absolute zero is the temperature at which atomic and molecular translational motion ceases. Temperatures above this are given in degrees Kelvin (K). The absolute zero is 273.15 K below the freezing point of pure water at 1 atm pressure, i.e., 0°C is 273.15 K.

Activity coefficient Activity coefficients are parameters that relate the thermodynamic activities of dissolved molecules to their physical concentrations. When molecules or ions in solution are far enough apart so that no mutual forces are exerted on one another, they follow the thermodynamic laws of ideal solutions. In a solution, the chemical potential of species i , μ_i , is given by

$$\mu_i = (\partial G / \partial n_i),$$

where n_i is the number of moles of molecular species i and G is the total free energy of the solution. That is, μ_i is the partial molal free energy. In an ideal

solution

$$\mu_i = \mu_{i,0} + RT \ln x_i,$$

where x_i is the mole fraction of the species i , R is the universal gas constant (see **physical constants**), and T is the **absolute temperature**. The term $\mu_{i,0}$ is the chemical potential of the species under a defined standard condition. In a non-ideal solution (e.g., a concentrated electrolytic solution), a function, the activity, must be introduced in place of x_i to account for deviations from ideality. The activity, a_i , is defined by

$$a_i = f_i x_i,$$

where f_i is the activity coefficient. Then,

$$\mu_i = \mu_{i,0} + RT \ln x_i + RT \ln f_i = \mu_{i,0} + RT \ln a_i.$$

The activity coefficient is, in general, a function of the composition of the solution. Theoretical treatments of nonideal solutions, such as the **Debye-Hückel theory**, attempt to predict experimentally determined activity coefficients.

Boltzmann's distribution law This distribution law stems from statistical mechanics (i.e., the theory of the behavior of large numbers of mass-particles in gas, liquid, or solid phases). It describes the population distribution of allowed energy states for the mass-particles. As derived by Ludwig Boltzmann in 1886, the law states that if N_0 is the number of molecules in any given state, the number N in a state whose potential energy is ΔE above that of the given state is

$$N = N_0 \exp(-\Delta E/kT),$$

where k is Boltzmann's constant ($k = R/N$, where N is Avogadro's number). Although derived without regard to quantum mechanics (which had not been discovered in 1886; see **Quantum mechanics: electrons and photons**), the law applies to populations of molecules with quantized energy states.

Brownian motion This term refers to random motion of small particles suspended in a fluid. The motions are caused by statistical fluctuations in the net momentum exchange between the molecules of the fluid and the suspended particles when they collide with each other. Another term used to describe a random process such as Brownian motion is "stochastic."

Chirality If a molecule cannot be superimposed on its mirror image it is said to be *chiral*, a word that is derived from the Greek word for hand. Hands are chiral objects because right and left hands have a nonsuperimposable mirror image relationship. Molecules that have the same molecular bonding skeleton, but differ in the absolute arrangement of atoms in space, are called *stereoisomers*. Stereoisomers that are related as object and nonsuperimposable mirror image are termed *enantiomers*. The fact that an enantiomeric stereoisomer can exist in a variety of conformations (three-dimensional geometries) does not destroy the enantiomeric relationship. That is, it is not possible to interconvert distinct enantiomeric isomers by free rotations about single bonds. A fundamental principle of organic chemistry is that a tetrahedral carbon atom with four chemically distinct substituents (an "asymmetric" carbon atom) within a molecule will result in enantiomeric forms. Each such carbon atom

forms a separate chiral center within the whole molecule. A section of a macromolecule, or a whole macromolecule, can have chirality in the same sense that right-handed and left-handed screw threads have overall chirality. The basis of the contribution of optical rotation and circular dichroism phenomena to biophysical measurements lies in the chirality of macromolecules in different conformations.

Debye–Hückel theory This widely accepted theory predicts the thermodynamic properties (i.e., parameters such as **activity coefficients**) of *dilute* electrolytic solutions. The theory begins with recognition that Coulomb's law gives the electrostatic force acting between individual pairs of electrically charged particles suspended in a medium as

$$F = e_1 e_2 / \epsilon r^2,$$

where e_1 and e_2 are their respective electric charges, ϵ is the **dielectric constant** of the medium, and r is the distance between the charges. This force is called a long-range force because it decreases only as $1/r^2$ (a very short-range force would decrease, for example, as $1/r^6$ to $1/r^{12}$). In an electrolytic solution containing highly dissociated salts, even though the ions are well separated, the total effect of all the other ions present on the potential energy of a given individual ion is large because of the long range of coulombic forces. The influence may be calculated for very dilute solutions of simple electrolytes by an approximation developed by P. Debye and E. Hückel in 1923. The central result of the approximation is that the electrostatic energy is a simple addition to the free energy of the collection of ions. If we consider various kinds of positive and negative ions of charge $z_i e$ (where $z_i = \pm 1, \pm 2$, etc., depending on the charge on the ion) at concentrations n_i and e is the electrical charge of the electron in coulombs (C), the approximation predicts that the electrostatic energy per mole of ions is

$$E = -\frac{1}{2} z_i^2 e^2 KN / \epsilon,$$

where N is Avogadro's number and K is given by

$$K = \left[8\pi e^2 / (\epsilon kT) \frac{1}{2} \sum n_i z_i^2 \right]^{1/2}.$$

In this expression for K , the term $\frac{1}{2} \sum c_i z_i^2$ is known as the ionic strength and is usually given the symbol I , and c_i is expressed in molarity of ions of a given species. Using this additional electrostatic energy formulation, various thermodynamic relations can be derived. For example, an important thermodynamic parameter is the activity coefficient f_i for an electrolytic solution. The natural logarithm of the activity coefficient of an ion *in a dilute solution*, on the basis of Debye–Hückel, is given by

$$\ln f_i = -z_i^2 e^2 KN / 2\epsilon RT.$$