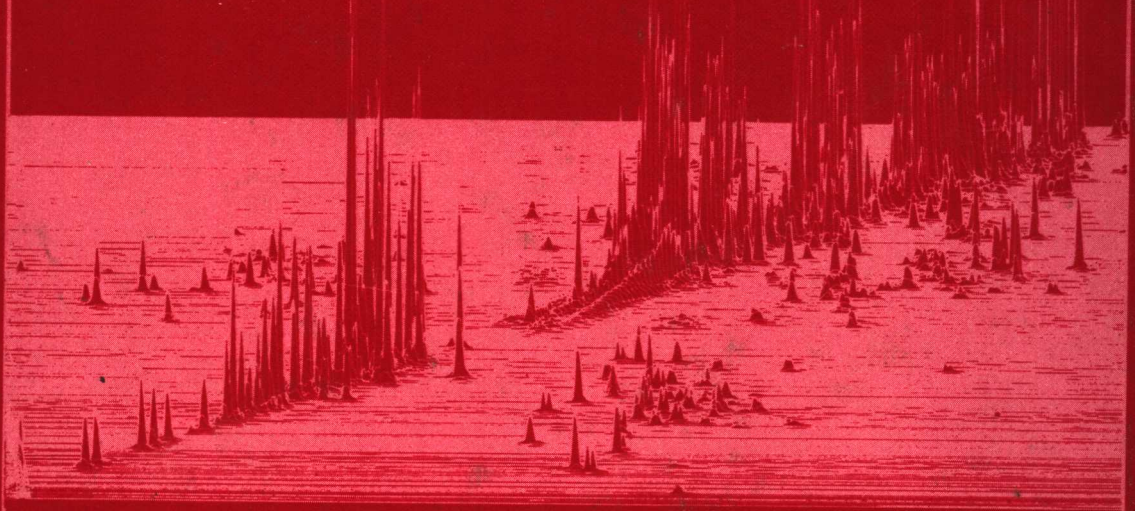
STRUCTURE AND DYNAMICS OF NUCLEIC ACIDS, PROTEINS, AND MEMBRANES



Edited by
E. Clementi
and
S. Chin

Structure and Dynamics of Nucleic Acids, Proteins, and Membranes

Edited by

E. Clementi

and

S. Chin

International Business Machines Corporation
Kingston, New York

PLENUM PRESS • NEW YORK AND LONDON

Library of Congress Cataloging in Publication Data

International Symposium on Structure and Dynamics of Nucleic Acids, Proteins, and Membranes (1986: Riva, Italy)
Structure and dynamics of nucleic acids, proteins, and membranes.

"Proceedings of the International Symposium on Structure and Dyanmics of Nucleic Acids, Proteins, and Membranes, held August 31–September 5, 1986, in Riva del Garda, Italy, under the sponsorship of the National Foundation for Cancer Research and the International Business Machines Corporation"—T.p. verso.
Includes bibliographies and index.

1. Molecular dynamics—Congresses. 2. Molecular structure—Congresses. 3. Proteins—Congresses. 4. Nucleic acids—Congresses. 5. Membranes (Biology)—Congresses. I. Clementi, Enrico. II. Chin, S. (Steven), 1957- . III. National Foundation for Cancer Research. IV. International Business Machines Corporation. V. Title. [DNLM: 1. Membrane Proteins—congresses. 2. Nucleic Acid Conformation—congresses. 3. Protein Conformation—congresses. 4. Structure–Activity Relationship—congresses. QU 55 I6738s 1986]
QP517.M65I58 1986 574.87'328 87-2366
ISBN 0-306-42553-X

Proceedings of the International Symposium on Structure and Dynamics of Nucleic Acids, Proteins, and Membranes, held August 31–September 5, 1986, in Riva del Garda, Italy, under the sponsorship of the National Foundation for Cancer Research and the International Business Machines Corporation

A Division of Plenum Publishing Corporation
233 Spring Street, New York, N.Y. 10013

Printed in the United States of America

Preface

This volume collects a number of the invited lectures and a few selected contributions presented at the *International Symposium on Structure and Dynamics of Nucleic Acids, Proteins and Membranes* held August 31st through September 5th, 1986, in Riva del Garda, Italy. The title of the conference as well as a number of the topics covered represent a continuation of two previous conferences, the first held in 1982 at the University of California in San Diego, and the second in 1984 in Rome at the Accademia dei Lincei. These two earlier conferences have been documented in *Structure and Dynamics: Nucleic Acids and Proteins,* edited by E. Clementi and R.H. Sarma, Adenine Press, New York, 1983, and *Structure and Motion: Membranes, Nucleic Acids and Proteins,* edited by E. Clementi, G. Corongiu, M.H. Sarma and R.H. Sarma, Adenine Press, New York, 1985.

At this conference in Riva del Garda we were very hesitant to keep the name of the conference the same as the two previous ones. Indeed, a number of topics discussed in this conference were not included in the previous ones and even the emphasis of this gathering is only partly reflected in the conference title. An alternative title would have been *Structure and Dynamics of Nucleic Acids, Proteins, and Higher Functions,* or, possibly, "higher components" rather than "higher functions." However, since these titles would have been somewhat ambiguous, in the end we decided to keep the original conference title — in deference to tradition — and to amend the ambiguities in this introduction.

The International Symposium in Riva del Garda, as well as this volume, are structured into several subfields ranging from relatively simple, to less simple, to relatively complex, to even more complex. To be correct, the so-called "relatively simple" is, however, quite complicated and it deals with structure of macromolecules. The chapters by Basosi et al., DeSantis et al., Middendorf et al., Poltev et al., Scheraga, and Wuthrich report about studies mainly on the structure of proteins and DNA, the grand foundation of macromolecules. Indeed, it is well known, especially to chemists, that structure is the first and necessary characterization of macromolecules. With the help of conformational analyses and simulations, DeSantis et al. attempted to explain DNA's supercoiling; Scheraga ventured as far as to predict — even if preliminarily — the structure of an interferon; while Wuthrich's mastery in obtaining NMR data opens the door to structural studies of proteins in solutions and offers detailed suggestions to molecular dynamicists. We then move in towards the dynamics of macromolecules with the papers by Alpert, Brunori et al., Dobson and Evans, Gratton et al., Karplus and Parak et al. It is well known that protein dynamics range from femto-seconds to hours, thus covering many time-scales. The many orders of time-scales represent a most difficult challenge for molecular dynamics computer simulations; in this context the contrib-

ution by Gratton et al. becomes critically important and appears to call for either a reinterpretation or an extension of the time-scales for the motion of some residues previously obtained from theory. Dynamics, of course, is a most fundamental and vital representation which holds the key to our understanding of the *function* of macromolecules. In this context we recall also the contribution by Sordo et al. dealing with *ab initio* interaction potentials, an alternative starting point for simulations of structural and dynamical studies, and the very interesting and appealing approach by Frauenfelder attempting to explain, with a relatively simple physical model — *a spinglass* — a very complex biological system — *a protein.*

But these macromolecules do not exist in a vacuum and interact with each other in the biological environment in which they exist; this observation brings out the importance of the study of the solvation of macromolecules, which is covered by the contributions of Careri and Rupley, Clementi et al., Lindsay, and Swamy and Clementi.

We should point out that Careri's presentation at Riva del Garda was notably broader than the one published here; by stressing "processes" and "correlations of events," rather than "structure and dynamics," one likely can interpret more deeply molecular and cellular biology.

Macromolecules — their structural interpretation and their dynamical characterization, and the environmental effects — are certainly three important and classical chapters of molecular biology. In this conference, however, we wanted to extend our learning and so we also included contributions from histology and modeling of cellular membranes; the chapters by Conti and Sackmann et al. deal with the latter.

It is self-evident that neurons are cells with very specific functions and are built up as a very complex organization of macromolecules; thus studies on the relationships between all the above chemical systems are an accepted viewpoint in science today.

However, in Riva del Garda we had sessions on artificial vision, pattern and voice recognition, artificial intelligence, present-day supercomputers, and even *sixth generation computers.* The chapters of Clementi et al., Mingolla and Grossberg, Reeke, and Schulten et al. are representative of these topics, possibly unrelated to biology, biophysics and physiology, would we accept a rather old fashioned — but still alive — vision of natural sciences. Our motivation for the inclusion of these modern topics of information sciences lies in the *determination to go much further* at this meeting and to link the most complex manmade machine with one of the most complex structures emerged from the evolutionary process; therefore, we asked what the components of the *computer* might have in common with the components of the *brain.*

The panel discussion did represent a very important moment of the symposium (it is a great pity that considerable portions of this discussion are not reported because of technical difficulties with the recording device). Berni Alder was a magnificent "moderator." Now, one might ask why we want to ask so many complicated and challenging questions at a single meeting. The answer is very simple: we have

reached a stage in our civilization where technology is being pushed to the very limit, where we are taking more and more risks because of our ignorance and greed, but also enthusiasm and creative curiousity. We are living in an era which, on one side, is filled with enormous excitement, new learning and discoveries, but, on the other side, might be sowing the seeds of great dangers for the immediate and/or long-range future. *Thus, we need to understand not only at a deeper level, but also at a broader level.* Interdisciplinarity is *not* a luxury — it is a need and, possibly, a *necessity*: to think about more and more powerful engines for computations and simulations, to attempt to realize pragmatically such machines, to hope to learn new ways of interconnecting machine components from physiological research on neuronal networks, is clearly only a hope, but a *rational* one. There might be an element of naivete in our goal; maybe a wise person might characterize our efforts as somewhat *childish.* Yet, we are only at the beginning in an existential need for understanding, discussing and building what might tomorrow become a truly thinking machine. When confronted with novel challenges, namely at every "beginning," we — man — will always be "adolescents" and exhibit "childish" trends, but hopefully on higher and higher evolutionary levels.

In Riva del Garda we have placed "seeds" for the sixth generation computers; let us gather once more in 1988 and reconsider where we shall be.

Steven Chin
IBM Corporation
Kingston, New York

Enrico Clementi
IBM Corporation
Kingston, New York

Contents

Part I. STRUCTURE OF MACROMOLECULES

Conformational Analysis of Polypeptides and Proteins for the Study of Protein Folding, Molecular Recognition, and Molecular Design 1
H.A. Scheraga, *Cornell University, Ithaca, USA*

Conformation of Non-Crystalline Proteins Viewed by NMR 21
K. Wüthrich, *ETH, Zurich, Switzerland*

Structures and Superstructures in Periodical Polynucleotides 31
P. DeSantis, S. Morosetti, A. Palleschi and M. Savino, *Università di Roma, Roma, Italy*

Conformational Peculiarities and Biological Role of Some Nucleotide Sequences 51
V.I. Poltev, A.V. Teplukhin and V.P. Chuprina, *USSR Academy of Sciences, Pushchino, USSR*

Multifrequency ESR of an Isotopically Enriched Copper System in the Immobilized Phase: A Monte Carlo Approach 69
R. Basosi, M. Pasenkiewicz-Gierula, W. Froncisz, W.E. Antholine, A. Jesmanowicz and J.S. Hyde, *University of Siena, Italy, Jagiellonian University, Poland, and National Biomedical ESR Center, Milwaukee, USA*

Neutron Scattering from Agarose Gels 75
F. Cavatorta, A. Deriu and H.D. Middendorf, *University of Parma, Italy, and University of Oxford, U.K.*

Non-Empirical Pair Potentials for the Interaction Between Amino Acids 89
J.A. Sordo, M. Probst, S. Chin, G. Corongiu and E. Clementi, *IBM Corporation, Kingston, USA*

Part II. DYNAMICS OF MACROMOLECULES

Molecular Dynamics of Proteins 113
M. Karplus, *Harvard University, Cambridge, USA*

Proton NMR Studies of Protein Dynamics and Folding: Applications of Magnetization Transfer NMR 127
C.M. Dobson and P.A. Evans, *University of Oxford, U.K.*

Distributions and Fluctuations of Protein Structures Investigated by X-Ray Analysis and Mössbauer Spectroscopy 139
F. Parak, M. Fischer, E. Graffweg and H. Formanek, *Wilhelms Universität, Münster, and Botanisches Institut der Universität München, München, FRG*

Rotational Motions of Tryptophan and Tyrosine Residues in Proteins 149
E. Gratton, J.R. Alcala and G. Marriott, *University of Illinois at Urbana-Champaign, Urbana, USA*

Protein Fluctuations and Hemeprotein Affinity for Ligand 153
B. Alpert, *Université Paris VII, Paris, France*

Cytochrome Oxidase in Energy Transduction 161
M. Brunori, P. Sarti, G. Antonini, F. Malatesta and M.T. Wilson, *University of Rome La Sapienza, University of Rome Tor Vergata, Rome, Italy, and University of Essex, Colchester, UK*

Proteins and Glasses 169
H. Frauenfelder, *University of Illinois at Urbana-Champaign, Urbana, USA*

Part III. SOLVATION OF MACROMOLECULES

Global Ab Initio Simulations: Study of a Liquid as an Example 179
E. Clementi, G.C. Lie, L. Hannon, D.C. Rapaport and M. Wojcik, *IBM Corporation, Kingston, USA, and National Foundation for Cancer Research, Bethesda, USA*

Proton Conductivity of Hydrated Lysozyme Powders, Considered Within the Framework of Percolation Theory 213
G. Careri and J.A. Rupley, *Universita di Roma I, Roma, Italy, and University of Arizona, Tucson, USA*

Molecular Dynamics Simulations of Biomolecules in Water 219
K.N. Swamy and E. Clementi, *IBM Corporation, Kingston, USA, and National Foundation for Cancer Research, Bethesda, USA*

Low Frequency Coherent Vibrations of DNA: The Role of the Hydration Shell and Phosphate-Phosphate Interactions 239
S.M. Lindsay, *Arizona State University, Tempe, USA*

Part IV. COMPONENTS OF LARGER SYSTEMS

Elasticity, Structure and Dynamics of Cell Plasma Membrane and Biological Functions 251
E. Sackmann, H.P. Duwe, K. Zeman and A. Zilker, *Technische Universität München, Garching, FRG*

Neuronal Signaling. A Simple Thermodynamic Process Involving Complex Membrane Proteins 269
F. Conti, *Istituto di Cibernetica e Biofisica, CNR, Genova, Italy*

Panel Discussions: Session 1. Understanding Brain Mechanisms and the Challenge of Artificial Intelligence Machines: An Interdisciplinary Approach; Session 2. Sixth Generation Computers: A Sketch 281

Physicists Explore Human and Artificial Intelligence 301
J. Buhmann, R. Divko, H. Ritter and K. Schulten, *Technische Universität München, Garching, FGR*

Recognition Automata Based on Selective Neural Networks 329
G.N. Reeke, Jr. and G.M. Edelman, *The Rockefeller University, New York, USA*

A Neural Theory of Preattentive Visual Information Processing: Emergent Segmentation, Cooperative-Competitive Computation, and Parallel Memory Storage 355
S. Grossberg and E. Mingolla, *Boston University, Boston, USA*

Large-Scale Computations on a Scalar, Vector and Parallel "Supercomputer" 403
E. Clementi, J. Detrich, S. Chin, G. Corongiu, D. Folsom, D. Logan, R. Caltabiano, A. Carnevali, J. Helin, M. Russo, A. Gnudi and P. Palmidese, *IBM Corporation, Kingston, USA*

Index 451

Conformational Analysis of Polypeptides and Proteins for the Study of Protein Folding, Molecular Recognition, and Molecular Design*

Harold A. Scheraga
Baker Laboratory of Chemistry
Cornell University
Ithaca, New York 14853-1301, USA

Abstract

Conformational energy calculations provide an understanding as to how interatomic interactions lead to the three-dimensional structures of polypeptides and proteins, and how these molecules interact with other molecules. Illustrative results of such calculations pertain to model systems (α-helices and β-sheets, and interactions between them), to various open-chain and cyclic peptides, to fibrous proteins, to globular proteins, and to enzyme-substrate complexes. In most cases, the validity of the computations is established by experimental tests of the predicted structures.

Key words: Internal interactions, ribonuclease, protein folding, conformational energy calculations, structural elements of proteins, oligopeptides, fibrous proteins, homologous proteins, globular proteins, enzyme-substrate complexes

1. Introduction

Twenty-five years ago, we began to develop computational methods[1] for the conformational analysis of polypeptides and proteins to study protein folding, protein structure, and interactions between proteins and other molecules (molecular recognition), and to suggest loci for site-specific mutations in natural proteins to modify their structure and reactivity (molecular design). The motivation for this development derived from our early experimental work on the structure of bovine pancreatic ribonuclease.[2] Before the X-ray structure of ribonuclease was known, we used solution physical chemical methods to identify three non-covalent tyrosyl . . . aspartate interactions,[2] whose locations were confirmed by the subsequently-determined X-ray structure[3] (Fig. 1). These, together with the known locations of

* This paper first appeared in the *Israel Journal of Chemistry*, Vol. 27, 1986.

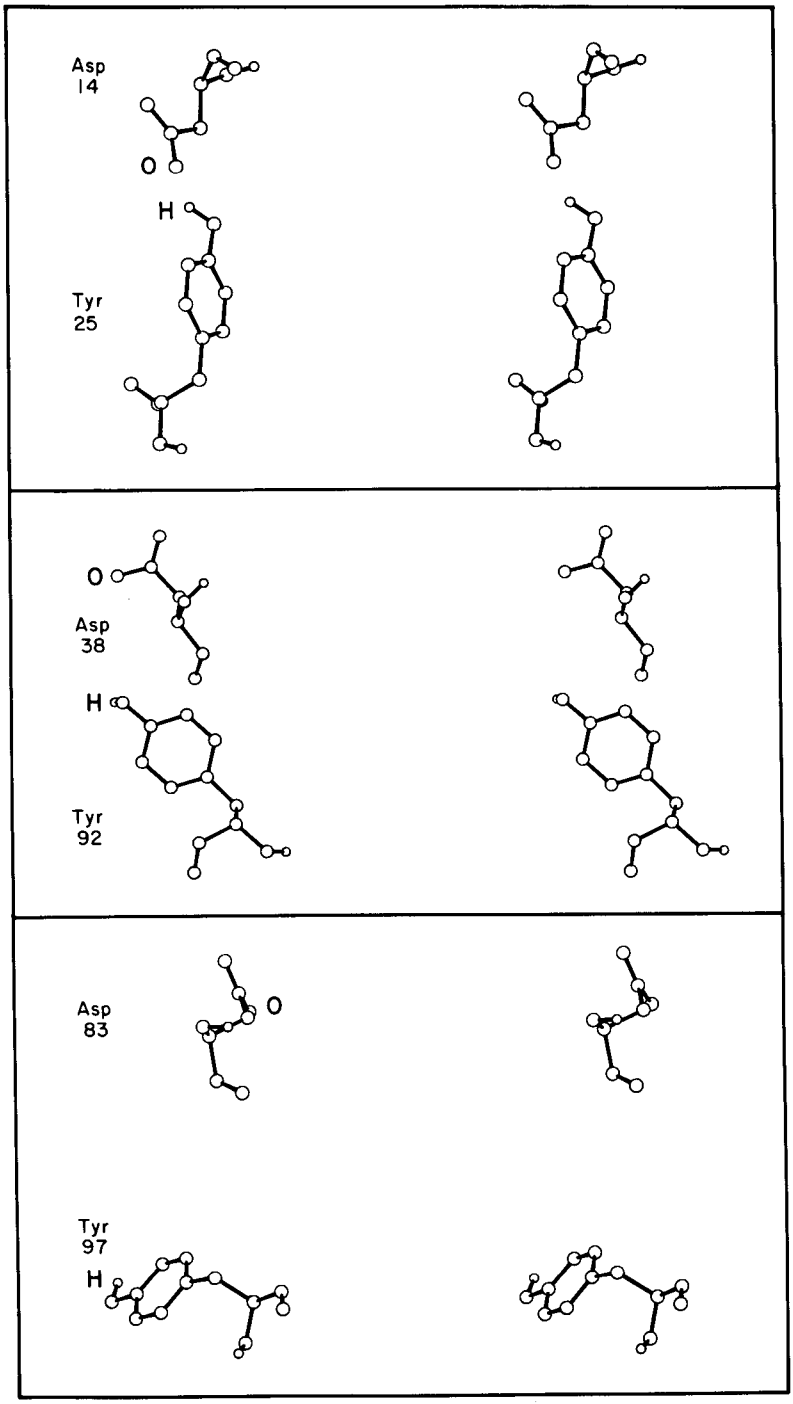


Figure 1. Three Tyr . . . Asp interactions deduced from solution physical chemical experiments.[2] The drawing is based on the subsequently-determined X-ray coordinates of Wlodawer et al.[3]

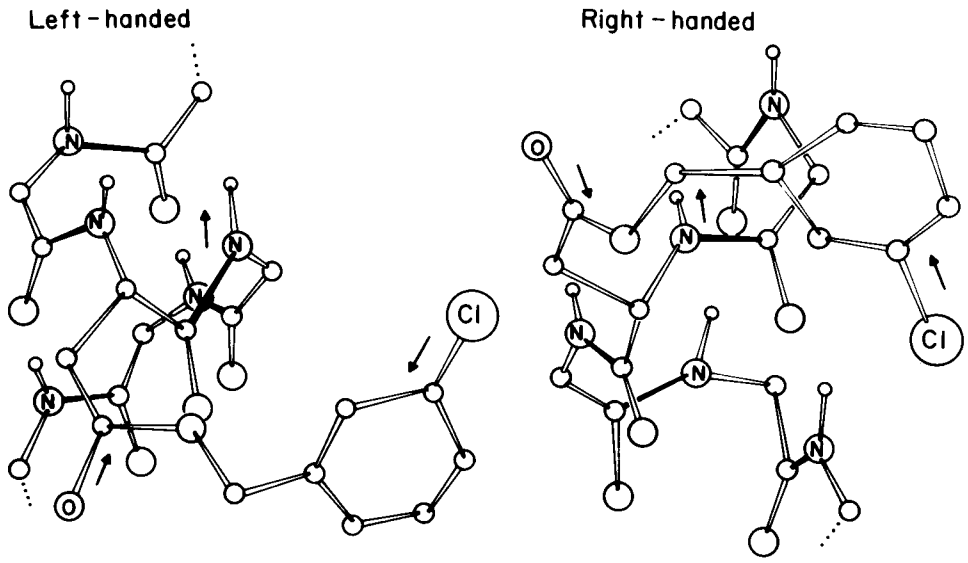


Figure 2. Orientation of the side chains of the lowest-energy left- and right-handed α-helices of poly (m-Cl-benzyl-L-aspartate).[22] The arrows represent the directions of the C-Cl, ester, and amide dipoles, respectively. The dipole-dipole interactions are more favorable in the left-handed form.

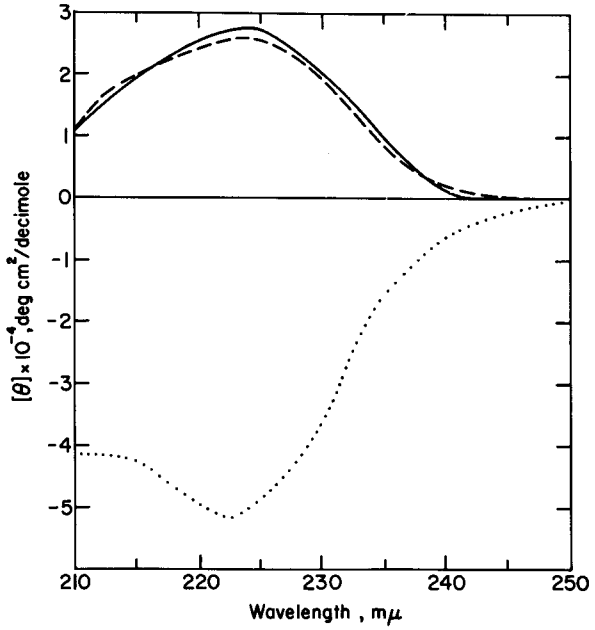


Figure 3. Circular dichroism spectra[32] of poly(β-benzyl-L-aspartates) in dioxane at 25°C. The symbols (——), (– – –), and (. . .) correspond to o-, m-, and p-Cl-benzyl derivatives, respectively.

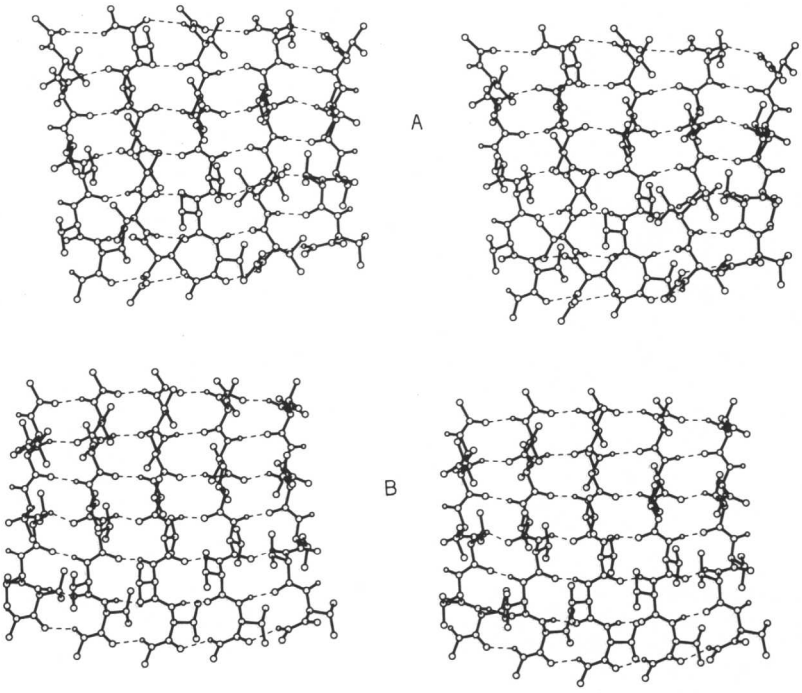


Figure 4. Stereo drawings of the minimum-energy β-sheets with five CH_3CO-(L-Val)$_6$-$NHCH_3$ chains.[24] (A) Antiparallel structure. (B) Parallel structure.

the four disulfide bonds, and the proximity of His 12, Lys 41 and His 119 in the active site of this enzyme, provided distance information that could serve as useful constraints to determine the three-dimensional structures of proteins by means of conformational energy calculations.

Initially,[1,4,5] only hard-sphere potentials were used in the computations, but subsequently more complete empirical energy functions were developed both in our laboratory[6-10] and elsewhere.[11-14] Our current main program is ECEPP/2, Empirical Conformational Energy Program for Peptides.[9,10] Procedures have also been introduced to take hydration and entropy effects into account, and to carry out energy minimization, Monte Carlo and molecular dynamics computations in a multidimensional space; these procedures have been reviewed on numerous occasions.[8,15-21]

While improvements in potential functions can be expected, the present ones are sufficiently accurate to provide agreement between various computational and experimental results, as will be shown here. The main obstacle to further progress is the multiple-minima problem, and much effort is being devoted to surmount this difficulty. In fact, this problem has already been solved for small open-chain and cyclic peptides and for fibrous proteins such as collagen, and progress is being made in the area of globular proteins.

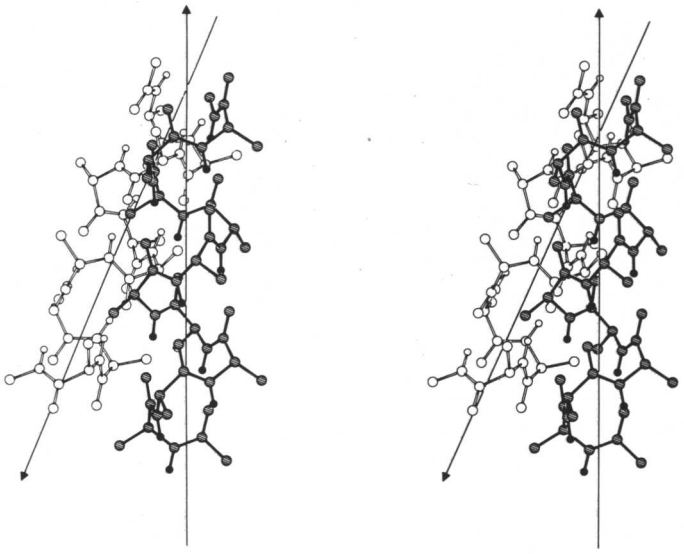


Figure 5. Stereo drawing of two CH_3CO-(L-Ala)$_{10}$-$NHCH_3$ α-helices in the lowest-energy packing arrangement.[27] The helix axes are indicated by arrows, with the head of the arrow pointing in the direction of the C-terminals of each helix.

This article will describe some of the conformational problems that have been elucidated by this computational methodology, and will summarize some of the current efforts being made to overcome the multi-minima problem for globular proteins. We shall describe first the results of computations on model systems of increasing complexity, and then consider the conformations of small open-chain and cyclic

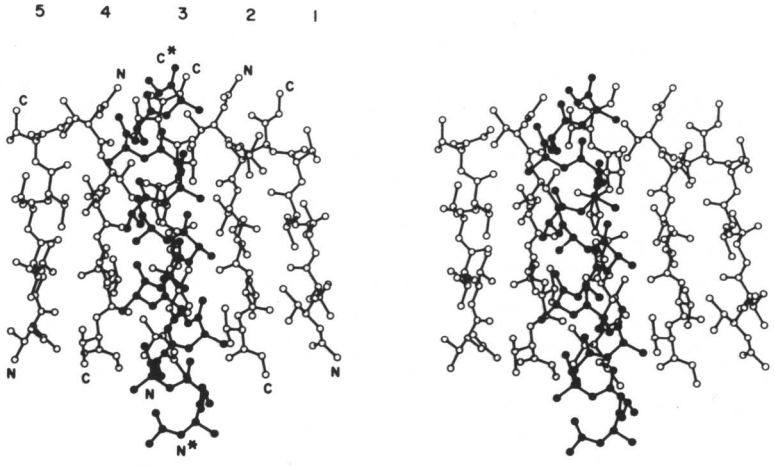


Figure 6. Stereo drawing of the lowest-energy packing arrangement of a CH_3CO-(L-Ala)$_{16}$-$NHCH_3$ α-helix and an antiparallel CH_3CO-(L-Val)$_6$-$NHCH_3$ β-sheet.[29]

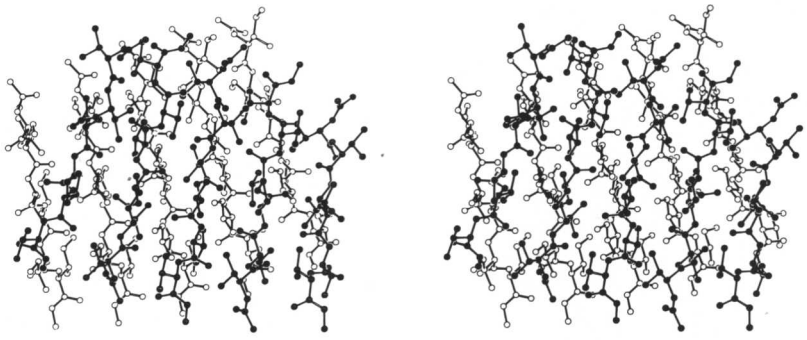


Figure 7. Stereo drawing of the lowest-energy packing arrangement of a CH_3CO-(L-Ile)$_6$-$NHCH_3$ parallel-chain β-sheet (open atoms) and a CH_3CO-(L-Val)$_6$-$NHCH_3$ antiparallel-chain β-sheet (filled atoms).[30]

peptides, fibrous proteins, and globular proteins. Further details can be found in several recent reviews.[18-21]

2. *Model Systems*

Conformational energy calculations have demonstrated how interatomic interactions lead to the preferred twists of α-helices[7,22] and β-sheets,[23-26] and to the preferred modes of packing of α-helices with α-helices,[25,27,28] α-helices with β-sheets,[29] β-sheets with β-sheets,[30] and pairs of triple helices of collagen with each other.[31] Some examples are provided in Figs. 2–8. Figure 2 illustrates how side chain-backbone dipole-dipole interactions play a dominant role in leading to a left-handed helix in poly(m-Cl-benzyl-L-aspartate).[22] The predicted[22] left-handedness of

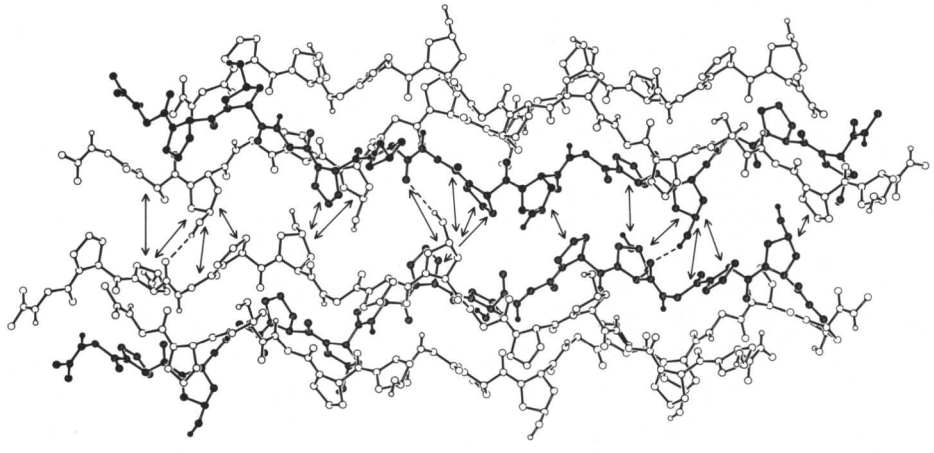


Figure 8. Computed lowest-energy packing arrangement[31] of two $[CH_3CO$-(Gly-Pro-Hyp)$_5$-$NHCH_3]_3$ triple helices, showing the near-parallel alignment of the two triple helices and O-H . . . O=C hydrogen bonds (dashed lines) between the triple helices. The arrows indicate residues that are in contact between the two triple helices.

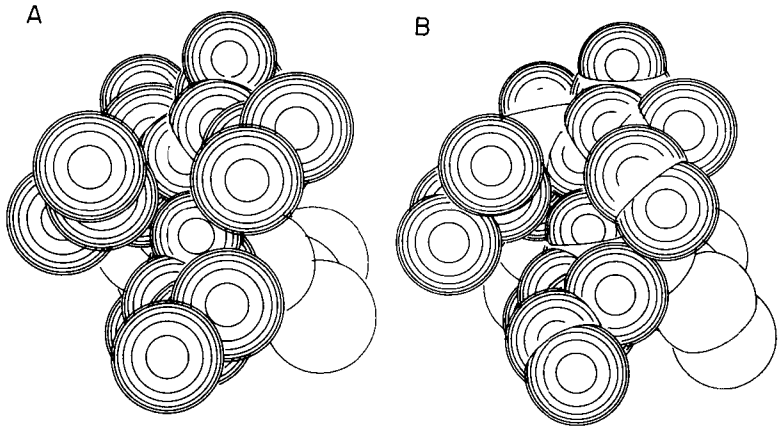


Figure 9. Space-filling models[36] of α-helical poly(L-valine) (A) and poly(L-isoleucine) (B).

the o- and m-derivatives, and the right-handedness of the p-derivative, were subsequently verified by circular dichroism measurements,[32] as indicated in Fig. 3. Figure 4 shows the computed right-handed twist of poly-L-valine β-sheets.[24] The lowest-energy packing arrangements for α . . . α, α . . . β, β . . . β pairs, and pairs of triple-helical collagen structures are illustrated in Figs. 5–8, respectively.

Conformational transitions have also been treated by this methodology, e.g. the interconversion of the cis and trans forms of poly(L-proline),[33] of the α- and

Table I. Comparison of Lowest-energy Structures Obtained by Two Different Methods[a]

	SMAPPS	Build-up plus Energy Minimization
	Backbone Dihedral Angles (DEG)	
Tyr ϕ	−86	−87
ψ	154	155
Gly ϕ	−155	−157
ψ	95	96
Gly ϕ	71	71
ψ	−90	−91
Phe ϕ	−90	−91
ψ	−40	−38
Met ϕ	−165	−165
ψ	−50	−48
Energy (kcal/mol)	−10.64	−10.66

[a]In this computation, the side chains were constrained to have the conformations of Fig. 12. More recently, this constraint has been eliminated (G.H. Paine and H.A. Scheraga, unpublished results).