Computational Methods For Protein Structure Prediction And Modeling Volume 1 Basic Characterization Biological And Medical Physics Biomedical Engineering

Protein Structure deals with the chemistry and physics of biologically important molecules—the proteins—particularly the determination of the structure of various proteins, their thermodynamics, their kinetics, and the mechanisms of different reactions of individual proteins. The book approaches the study of protein structure in two ways: firstly, by determining the general features of protein structure, the overall size, and shape of the molecule; and secondly, by investigating the molecule internally along with the various aspects of the internal configuration of protein molecules. It describes in detail experimental methods for determining protein structure in solution, such as the hydrodynamic method, the thermodynamic optical method, and the electrochemical method. The book then explains the results of experiments carried out on insulin, lysozyme, and ribonuclease. The text notes that the experiments, carried out on native and denatured proteins as well as on derivatives prepared by chemical modification (e.g., by methylation, iodination, acetylation, etc.), can lead to greater understanding of secondary and tertiary structures of proteins of known sequence. The book is suitable for biochemists, micro-biologists, cellular researchers, or investigators involved in protein structure and other biological sciences related to muscle physiologists, geneticists, enzymologists, or immunologists.

The Latest Developments on the Role of Dynamics in Protein Functions Computational Approaches to Protein Dynamics: From Quantum to Coarse-Grained Methods presents modern biomolecular computational techniques that address protein flexibility/dynamics at all levels of theory. An international contingent of leading researchers in chemistry, physics, and biology show how these advanced methods provide insights into dynamic aspects of biochemical processes. A particular focus is on intrinsically disordered proteins (IDPs), which lack a well-defined three-dimensional structure and function as dynamic ensembles. The book covers a wide spectrum of dynamics, from electronic structure-based to coarse-grained techniques via multiscaling at different levels. After an introduction to dynamics and historical overview of basic methodologies, the book addresses the following issues: Is there a quantitative relationship between enzymatic catalysis and protein dynamics? Which are the functionally relevant motions of proteins? How can structural properties and partner recognition mechanisms of IDPs be simulated? How can we speed up molecular dynamics? How can we describe conformational ensembles by the synergistic effort of computations and experiments? While dynamics is now considered essential for interpreting protein action, it is not yet an integral component in establishing structure–function relationships of proteins. Helping to reshape this classical view in biochemistry, this groundbreaking book explores advances in computational methodology and contributes to the new, ensemble way of studying proteins.

This book comprises papers on diverse aspects of fuzzy logic, neural networks, and nature-inspired optimization meta-heuristics and their application in various areas such as intelligent control and robotics, pattern recognition, medical diagnosis, time series prediction and optimization of complex problems. The book is organized into seven main parts, each with a collection of papers on a similar subject. The first part presents new concepts and algorithms based on type-2 fuzzy logic for dynamic parameter adaptation in meta-heuristics. The second part discusses network theory and applications, and includes papers describing applications of neural networks in diverse areas, such as time series prediction and pattern recognition. The third part addresses the theory and practice of meta-heuristics in different areas of application, while the fourth part describes diverse fuzzy logic applications in the control area, which can be considered as intelligent controllers. The next two parts explore applications in areas, such as time series prediction, and pattern recognition and new optimization and evolutionary algorithms and their applications respectively. Lastly, the seventh part addresses the design and application of different hybrid intelligent systems.

An overview of algorithms important to computational structural biology that addresses such topics as NMR and design and analysis of proteins. Using the tools of information technology to understand the molecular machinery of the cell offers both challenges and opportunities to computational scientists. Over the past decade, novel algorithms have been developed both for analyzing biological data and for synthetic biology problems such as protein engineering. This book explains the algorithmic foundations and computational approaches underlying areas of structural biology including NMR (nuclear magnetic resonance); X-ray crystallography; and the design and analysis of proteins, peptides, and small molecules. Each chapter offers a concise overview of important concepts, focusing on a key topic in the field. Four chapters offer a short course in algorithmic and computational issues related to NMR structural biology, giving the reader a useful toolkit with which to approach the fascinating yet thorny computational problems in this area. A recurrent theme is understanding the interplay between biophysical experiments and computational algorithms. The text emphasizes the mathematical foundations of structural biology while maintaining a balance between algorithms and a nuanced understanding of experimental data. Three emerging areas, particularly fertile ground for research students, are highlighted: NMR methodology, design of proteins and other molecules, and the modeling of protein flexibility. The next generation of computational structural biologists will need training in geometric algorithms, provably good approximation algorithms, scientific computation, and an array of techniques for handling noise and uncertainty in combinatorial geometry and computational biophysics. This book is an essential guide for young scientists on their way to research success in this exciting field.

Understanding the synthesis, structures and functions of proteins draws vital attention in computational biology as proteins participate in virtually every cellular function in an organism. In appropriate environment, a protein folds spontaneously into unique three dimensional structure of minimum energy termed as native state. Protein Structure Prediction (PSP) refers to the computational approach of predicting protein tertiary structure from amino acid sequence. Protein synthesis, on the other hand, is a multi-step process where nuclear DNA is transcribed into protein-coding messenger RNA (mRNA), which is then translated into unique amino acid sequence. MicroRNAs (miRNAs) bind to target mRNAs through complementary base-pairing and regulate protein production by translational repression or target degradation. A miRNA can bind to another mRNA from a po¬tentially large mRNA pool and computational prediction of such target mRNA set is referred to as miRNA Target Prediction. -- Incomplete knowledge of folding mechanism, absence of an established perfect energy function, and apparently complex and irregular tertiary structure make the PSP problem ever so difficult, which encourages researchers adopting simplified lattice and energy models to ease the computa¬tional hardness of the problem so as to explain essential functional properties of proteins. This thesis aims at developing several stochastic optimisation approaches to ab initio PSP in triangular lattice models and comparing their relative efficacy. Triangular lattice models are chosen because of their ability to capture more compact folded structures. In search for a faster and efficient local search method, a new neighbourhood relation is developed that is shown complete and efficient in finding

minimum energy structures when incorporated into tabu search and logarithmic simulated annealing algorithm.

Computational biology is a rapidly expanding field, and the number and variety of computational methods used for DNA and protein sequence analysis is growing every day. These algorithms are extremely valuable to biotechnology companies and to researchers and teachers in universities. This book explains the latest computer technology for analyzing DNA, RNA, and protein sequences. Clear and easy to follow, designed specifically for the non-computer scientist, it will help biologists make better choices on which algorithm to use. New techniques and demonstrations are elucidated, as are state-of-the-art problems, and more advanced material on the latest algorithms. The primary audience for this volume are molecular biologists working either in biotechnology companies or academic research environments, individual researchers and the institutions they work for, and students. Any biologist who relies on computers should want this book. A secondary audience will be computer scientists developing techniques with applications in biology. An excellent reference for leading techniques, it will also help introduce computer scientists to the biology problems. This is an outstanding work which will be ideal for the increasing number of scientists moving into computational biology.

Abstract: Nearly all major processes in living cells are carried out by complex apparatus consisting of protein molecules. This thesis describes computational tools developed to help investigate two fundamental questions about proteins that underlie cell functions: how they interact with each other and form complex structures; and how they are expressed and modified in different cell states. In order to address the first question, several methods are developed to predict protein-protein complex structures. Protein interactions are energy driven processes. The prediction of protein complex structures is the search for the global minimum on the binding free-energy landscape. An approach is described that uses Van der Wools energy, desolvation energy and shape complementarity as the scoring functions and a five-dimensional fast Fourier transform algorithm to expedite the search. Two methods to screen and optimize the predicted protein complex structures are also introduced. They incorporate additional energy terms and clustering algorithms to provide more precise estimations of the binding free-energy. The same methods can also be used to predict hot spots, the mutations of which significantly alter the binding kinetics. To study the protein expression profiles, a two-step approach for protein identification using peptide mass fingerprinting data is developed. Peptide mass fingerprinting uses peptide masses determined by mass spectrometry to identify the peptides and subsequently, the proteins in the sample Peaks in the mass spectrum are associated with known peptide sequences in the database based on log-likelihood ratio test. A statistical algorithm is then used to identify proteins by comparing the probability of each protein's presence in the sample, given the peak assignments with the background probability. This method also discovers post-translational modifications in the identified proteins. The protein binding prediction program successfully predicts protein complex structures that closely resemble their native forms, as observed by x-ray crystallography or NMR. The refinements and hot spot predictions also give accurate and consistent results. The database search program that interprets mass spectrometry data is evaluated with artificial and experimental data. The program identifies proteins in the sample with high sensitivity and specificity. The results presented in this thesis demonstrate that computational methods help to better understand the structure and the composition of the protein machineries. All of the methods described herein have been implemented and made available for the research community over the Internet.

A Step-by-Step Guide to Describing Biomolecular Structure Computational and Visualization Techniques for Structural Bioinformatics Using Chimera shows how to perform computations with Python scripts in the Chimera environment. It focuses on the three core areas needed to study structural bioinformatics: biochemistry, mathematics, and computation. Understand Important Concepts of Structural Bioinformatics The book covers topics that deal primarily with protein structure and includes many exercises that are grounded in biological problems at the molecular level. The text encourages mathematical analysis by providing a firm foundation for computations. It analyzes numerous Python scripts for the Chimera environment, with the scripts and other material available on a supplementary website. Build Python Scripts to Extend the Capabilities of Chimera Through more than 60 exercises that involve the development of Python scripts, the book gives you concrete guidance on using the scripting capabilities of Chimera. You’ll gain experience in solving real problems as well as understand the various applications of linear algebra. You can also use the scripts as starting points for the development of similar applications and use classes from the StructBio toolkit for computations, such as structure overlap, data plotting, scenographics, and display of residue networks.

This tutorial was one of eight tutorials selected to be presented at the Third International Conference on Intelligent Systems for Molecular Biology which was held in the United Kingdom from July 16 to 19, 1995. The authors intend to review the state of the art in the experimental determination of protein 3D structure (focus on nuclear magnetic resonance), and in the theoretical prediction of protein function and of protein structure in 1D, 2D and 3D from sequence. All the atomic resolution structures determined so far have been derived from either X-ray crystallography (the majority so far) or Nuclear Magnetic Resonance (NMR) Spectroscopy (becoming increasingly more important). The authors briefly describe the physical methods behind both of these techniques; the major computational methods involved will be covered in some detail. They highlight parallels and differences between the methods, and also the current limitations. Special emphasis will be given to techniques which have application to ab initio structure prediction. Large scale sequencing techniques increase the gap between the number of known proteins sequences and that of known protein structures. They describe the scope and principles of methods that contribute successfully to closing that gap. Emphasis will be given on the specification of adequate testing procedures to validate such methods.

Proteins lie at the heart of almost all biological processes and have an incredibly wide range of activities. Central to the function of all proteins is their ability to adopt, stably or sometimes transiently, structures that allow for interaction with other molecules. An understanding of the structure of a protein can therefore lead us to a much improved picture of its molecular function. This realisation has been a prime motivation of recent Structural Genomics projects, involving large-scale experimental determination of protein structures, often those of proteins about which little is known of function. These initiatives have, in turn, stimulated the massive development of novel methods for prediction of protein function from structure. Since model structures may also take advantage of new function prediction algorithms, the first part of the book deals with the various ways in which protein structures may be predicted or inferred, including specific treatment of membrane and intrinsically disordered proteins. A detailed consideration of current structure-based function prediction methodologies forms the second part of this book, which concludes with two chapters, focusing specifically on case studies, designed to illustrate the real-world application of these methods. With bang up-to-date texts from world experts, and abundant links to publicly available resources, this book will be invaluable to anyone who studies proteins and the endlessly fascinating relationship between their structure and function.

This book provides a comprehensive overview of modern computer-based techniques for analyzing the structure, properties and dynamics of biomolecules and biomolecular processes. It is organized in four main parts; the first one deals with methodology of molecular simulations; the second one with applications of molecular simulations; the third one introduces bioinformatics methods and the use of experimental information in molecular simulations; the last part reports on selected applications of molecular quantum mechanics. This second edition has been thoroughly revised and updated to include the latest progresses made in the respective field of research.

Using a geometric perspective, Protein Geometry, Classification, Topology, and Symmetry reviews and analyzes the structural principals of proteins with the goal of revealing the underlying regularities in their construction. It also reviews computer methods for structure analysis and the automatic comparison and classification of these structures with an analysis of the statistical significance of comparing different shapes. Following an analysis of the current state of protein classification, the authors explore more abstract geometric and topological representations, including the occurrence of knotted topologies. The book concludes with a consideration of the origin of higher-level symmetries in protein structure. The authors focus on simple geometric methods that are deterministic rather than probabilistic and on the more abstract simplifications of protein structure that allow a better understanding of the overall fold of the structure. Most of the methods described in this book have corresponding computer programs. These can be found (as C source code) at the ftp site of the Division of Mathematical Biology at the National Institute for Medical Research. This collection of ideas contains pedagogical material that make it ideal for post-graduate courses as well as new ideas and results essential for researchers investigating protein structures.

Understanding sequence-structure relationships of proteins is a central theme of computational structural biology. To create accurate mapping between sequences and structures is a big computational challenge, because the inherent dynamics of protein molecules requires any structure to be seen as an ensemble containing a large number of structural states. In this thesis, I focus on developing new structural modeling methods representing two routes towards efficient sequence-structure mapping that are compatible with this ensemble view of structures. First, I will show that the relationships between the sequence and the structural ensemble of a protein can be revealed by breaking down the protein into constituent structural fragments, for which ensemble statistics can be obtained from the protein structure database. Second, sequence-structure relationships can be also extracted by combining explicit atomistic modeling of ensembles and statistical tools reducing the overall computational cost. Implications in structure prediction, mutational analysis, and design of protein-interaction modulators will be presented and discussed, showing the great promise held by these methods in further improving the state-of-the-art in a broad spectrum of applications in computational structural biology.

Molecular Modeling of Proteins, Second Edition provides a theoretical background of various methods available and enables non-specialists to apply methods to their problems by including updated chapters and new material not covered in the first edition. This detailed volume opens by featuring classical and advanced simulation methods as well as methods to set-up complex systems such as lipid membranes and membrane proteins and continues with chapters devoted to the simulation and analysis of conformational changes of proteins, computational methods for protein structure prediction, usage of experimental data in combination with computational techniques, as well as protein-ligand interactions, which are relevant in the drug design process. Written for the highly successful Methods in Molecular Biology series, chapters include thorough introductions, step-by-step instructions and notes on troubleshooting and avoiding common pitfalls. Update-to-date and authoritative, Molecular Modeling of Proteins, Second Edition aims to aid researchers in the physical, chemical and biosciences interested in utilizing this powerful technology.

Volume Two of this two-volume sequence presents a comprehensive overview of protein structure prediction methods and includes protein threading, De novo methods, applications to membrane proteins and protein complexes, structure-based drug design, as well as structure prediction as a systems problem. A series of appendices review the biological and chemical basics related to protein structure, computer science for structural informatics, and prerequisite mathematics and statistics.

This book discusses topics related to bioinformatics, statistics, and machine learning, presenting the latest research in various areas of bioinformatics. It also highlights the role of computing and machine learning in knowledge extraction from biological data, and how this knowledge can be applied in fields such as drug design, health supplements, gene therapy, proteomics and agriculture.

This book presents applications of bioinformatics tools that experimental research scientists use in "daily practice." Its interdisciplinary approach combines computational and experimental methods to solve scientific problems. The book begins with reviews of computational methods for protein sequence-structure-function analysis, followed by methods that use experimental data obtained in the laboratory to improve functional predictions.

This is a comprehensive introduction to Landau-Lifshitz equations and Landau-Lifshitz-Maxwell equations, beginning with the work by Yulin Zhou and Boling Guo in the early 1980s and including most of the work done by this Chinese group led by Zhou and Guo since. The book focuses on aspects such as the existence of weak solutions in multi dimensions, existence and uniqueness of smooth solutions in one dimension, relations with harmonic map heat flows, partial regularity and long time behaviors. The book is a valuable reference book for those who are interested in partial differential equations, geometric analysis and mathematical physics. It may also be used as an advanced textbook by graduate students in these fields.

This volume provides computational methods and reviews various aspects of computational studies of protein aggregation. Chapters discuss the relationship between protein misfolding and protein aggregation, methods of prediction of aggregation propensities of protein, peptides, protein structure, results of computer simulations of aggregation, and computational simulations focused on specific diseases such as Alzheimer’s, Parkinson’s, and preeclampsia. Written in the highly successful Methods in Molecular Biology series format, chapters include introductions to their respective topics, lists of the necessary materials and reagents, step-by-step, readily reproducible laboratory protocols, and tips on troubleshooting and avoiding known pitfalls. Authoritative and cutting-edge, Computer Simulations

of Aggregation of Proteins and Peptides aims to ensure successful results in the further study of this vital field.
This book discusses a broad range of basic and advanced topics in the field of protein structure, function, folding, flexibility, and dynamics. Starting with a basic introduction to protein purification, estimation, storage, and its effect on the protein structure, function, and dynamics, it also discusses various experimental and computational structure determination approaches; the importance of molecular interactions and water in protein stability, folding and dynamics; kinetic and thermodynamic parameters associated with protein-ligand binding; single molecule techniques and their applications in studying protein folding and aggregation; protein quality control; the role of amino acid sequence in protein aggregation; muscarinic acetylcholine receptors, antimuscarinic drugs, and their clinical significances. Further, the book explains the current understanding on the therapeutic importance of the enzyme dopamine beta hydroxylase; structural dynamics and motions in molecular motors; role of cathepsins in controlling degradation of extracellular matrix during disease states; and the important structure-function relationship of iron-binding proteins, ferritins. Overall, the book is an important guide and a comprehensive resource for understanding protein structure, function, dynamics, and interaction.
Volume One of this two-volume sequence focuses on the basic characterization of known protein structures, and structure prediction from protein sequence information. Eleven chapters survey of the field, covering key topics in modeling, force fields, classification, computational methods, and structure prediction. Each chapter is a self contained review covering definition of the problem and historical perspective; mathematical formulation; computational methods and algorithms; performance results; existing software; strengths, pitfalls, challenges, and future research.
This text offers in-depth perspectives on every aspect of protein structure identification, assessment, characterization, and utilization, for a clear understanding of the diversity of protein shapes, variations in protein function, and structure-based drug design. The authors cover numerous high-throughput technologies as well as computational methods to study protein structures and residues. A valuable reference, this book reflects current trends in the effort to solve new structures arising from genome initiatives, details methods to detect and identify errors in the prediction of protein structural models, and outlines challenges in the conversion of routine processes into high-throughput platforms.
The major goal of 'Expanding Frontiers in Polypeptide and Protein Structural Research' has been to bring the various avenues for the exploration of protein structures to a single forum. The idea of organizing the symposiwn was conceived by one of the editors, V. Renugopalakrislman, during the 9th International Biophysics Congress satellite symposium at Kibbutz Nof-Ginosar, Israel in 1987. It was originally supposed to dwell on 2D NMR and molecular dynamics of polypeptides and proteins. During the earlier part of the last decade, these two approaches began to emerge as powerful tools to probe protein structures at the atomic level in solution. The developments in molecular biology ushered in the capability to design polypeptides and proteins for specific application in science and technology. The emergence of 2D NMR and molecular dynamics was greatly facilitated by contemporary developments in molecular biology and protein engineering. Therefore an international symposiwn devoted exclusively to 2D NMR and molecular dynamics studies of proteins was felt necessary to bring two major approaches in a single forum. In addition to emphasis on 2D NMR and molecular dynamics simulation, the scope of the symposiwn included optical spectroscopy, protein design, and new horizons in protein structure. The symposiwn consisted of five plenary sessions devoted to NMR and optical spectroscopy as probes for protein structure, protein dynamics, computational methods in protein design, and new horizons in protein structure. In addition, five workshops in related areas, viz.
The aim this volume is to present the methods, challenges, software, and applications of this widespread and yet still evolving and maturing field. Computational Protein Design, the first book with this title, guides readers through computational protein design approaches, software and tailored solutions to specific case-study targets. Written in the highly successful Methods in Molecular Biology series format, chapters include introductions to their respective topics, step-by-step, readily reproducible laboratory protocols, and tips on troubleshooting and avoiding known pitfalls. Authoritative and cutting-edge, Computational Protein Design aims to ensure successful results in the further study of this vital field.
Proteins: Structure and Function is a comprehensive introduction to the study of proteins and their importance to modern biochemistry. Each chapter addresses the structure and function of proteins with a definitive theme designed to enhance student understanding. Opening with a brief historical overview of the subject the book moves on to discuss the ‘building blocks’ of proteins and their respective chemical and physical properties. Later chapters explore experimental and computational methods of comparing proteins, methods of protein purification and protein folding and stability. The latest developments in the field are included and key concepts introduced in a user-friendly way to ensure that students are able to grasp the essentials before moving on to more advanced study and analysis of proteins. An invaluable resource for students of Biochemistry, Molecular Biology, Medicine and Chemistry providing a modern approach to the subject of Proteins.
Since the first attempts to model proteins on a computer began almost thirty years ago, our understanding of protein structure and dynamics has dramatically increased. Spectroscopic measurement techniques continue to improve in resolution and sensitivity, allowing a wealth of information to be obtained with regard to the kinetics of protein folding and unfolding, and complementing the detailed structural picture of the folded state. Concurrently, algorithms, software, and computational hardware have progressed to the point where both structural and kinetic problems may be studied with a fair degree of realism. Despite these advances, many major challenges remain in understanding protein folding at both the conceptual and practical levels. Computational Methods for Protein Folding seeks to illuminate recent advances in computational modeling of protein folding in a way that will be useful to physicists, chemists, and chemical physicists. Covering a broad spectrum of computational methods and practices culled from a variety of research fields, the editors present a full range of models that, together, provide a thorough and current description of all aspects of protein folding. A valuable resource for both students and professionals in the field, the book will be of value both as a cutting-edge overview of existing information and as a catalyst for inspiring new studies. Computational Methods for Protein Folding is the 120th volume in the acclaimed series Advances in Chemical Physics, a compilation of scholarly works dedicated to the dissemination of contemporary advances in chemical physics, edited by Nobel Prize-winner Ilya Prigogine.
This volume presents a diverse collection of methodologies used to study various problems at the protein sequence and structure level. The chapters in this book look at issues ranging from broad concepts like protein space to specifics like antibody modeling. Topics include point mutations, gene duplication, de novo emergence of new genes, pairwise correlated mutations, ancestral protein reconstruction, homology modelling, protein stability and dynamics, and protein-protein interactions. The book also covers a wide range of computational approaches, including sequence and structure alignments, phylogenies, physics-based and mathematical approaches, machine learning, and more. Written in the highly successful Methods in Molecular Biology series format, chapters include introductions to their respective topics, lists of the necessary materials and prerequisites, step-by-step, readily reproducible computational protocols (using command line or graphical user interfaces, sometimes including computer code), and tips on troubleshooting and avoiding known pitfalls. Cutting-edge and authoritative, Computational Methods in Protein Evolution is a valuable resource that offers useful workflows and techniques that will help both novice and expert researchers working with proteins computationally.
This book covers elements of both the data-driven comparative modeling approach to structure prediction and also recent attempts to

simulate folding using explicit or simplified models. Despite the unsolved mystery of how a protein folds, advances are being made in predicting the interactions of proteins with other molecules. Also rapidly advancing are the methods for solving the inverse folding problem, the problem of finding a sequence to fit a structure. This book focuses on the various computational methods for prediction, their successes and their limitations, from the perspective of their most well known practitioners.

Computational Methods for Protein Structure Prediction and ModelingVolume 1: Basic CharacterizationSpringer Science & Business Media

Despite significant advancement being made during the recent past in predicting structure of proteins using computational methods, these techniques often cannot achieve sufficiently high level of accuracy to fully appreciate biological function and to serve as a reliable starting point for rational drug design efforts to develop novel therapeutics. Bringing these low-resolution models as close as possible to the native structure, called the protein structure refinement problem, however, has remained largely unsolved. Existing approaches in protein structure refinement suffer from two key challenges: (1) lack of consistency and (2) failure to produce meaningful degree of refinement. This thesis is composed of three major contributions. First, we propose a consistent and computationally efficient computational optimization protocol called 3Drefine. Next, we further improve the 3Drefine algorithm by developing an iterative version of the protocol, named i3Drefine. Finally, we present a novel conformation ensemble-based iterative refinement method, REFINEpro, aimed at producing pronounced degree of refinement. All of these methods were benchmarked in large-scale benchmark datasets and achieved consistent refinement in both global and local structural quality measures. In particular, i3Drefine was ranked as the best protein structure refinement server method in recent Critical Assessment of Protein Structure Prediction experiment. All of these methods are freely available to the scientific community in the form of software and web-servers.

The present invention provides a method utilizing primary amino acid sequence of a protein, energy minimization, molecular dynamics and protein vibrational modes to predict three-dimensional structure of a protein. The present invention also determines possible intermediates in the protein folding pathway. The present invention has important applications to the design of novel drugs as well as protein engineering. The present invention predicts the three-dimensional structure of a protein independent of size of the protein, overcoming a significant limitation in the prior art.