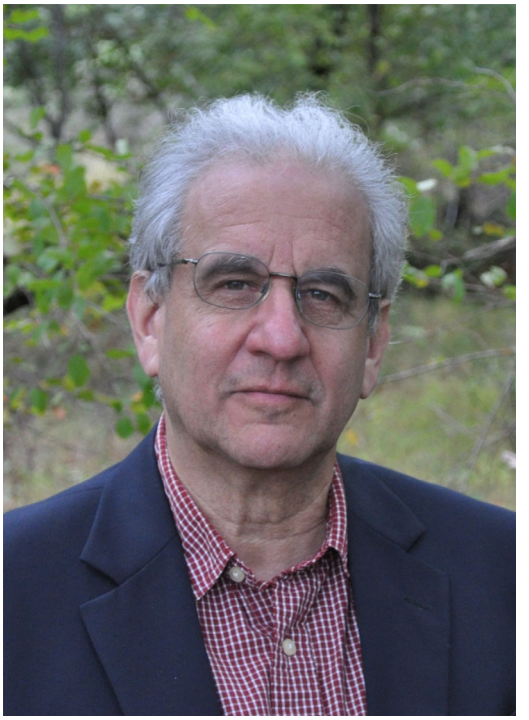
THE PROTEIN SOCIETY

EDITORIAL

It is our distinct pleasure to present this special issue of *Protein Science* in recognition of the significant and extensive contributions of Ronald M. Levy, the Laura H. Carnell Professor of Biophysics and Computational Biology and Professor of Chemistry, Physics, and Biology, and in celebration of his 65th birthday. With this issue if is evident that the breadth and scope of the contributions from Ron's colleagues and friends truly represent the impact he has had on the field throughout his career.

Ron was an undergraduate at Reed College in Portland, Oregon, where he graduated in 1970 with majors in Biology and Math. He did his Ph.D. and postdoctoral studies at Harvard University with Martin Karplus, 2013 Nobel Prize winner in Chemistry. Ron focused on Biophysics and played an important role in sculpting the landscape of computational biophysical modeling and simulation. Ron began his independent career at Rutgers University as an Assistant Professor of Chemistry in 1980. He remained at Rutgers throughout much of his career; advancing in academic rank and helping Rutgers build an impressive program at the interfaces of Biology, Mathematics, Physics and Chemistry as initially the Co-Director and then Director of the BioMaPS Institute for Quantitative Biology. In 2014 Ron moved his research program to Temple University, where he holds the positions noted above as well as directing the Center for Biophysics and Computational Biology.

Ron's research and publications (nearly 200) span a broad range of fundamental developments linking statistical mechanics to computational methodologies for studying biomacromolecules, important contributions to the theory of NMR applied to proteins, key applications to a plethora of important biological systems for which he and his co-workers have enlightened and informed us of the inner workings of complex biological phenomena. In short, Ron's scientific career, his methodological developments, applications, and publications to date have had a lasting impact on the field of theoretical and computational biophysics. It is these contributions that we celebrate with this special issue. The impressive response of Ron's colleagues, and the exceptional quality of the collected papers here well represent the impact Ron has had, and continues to have, on the field. In the following paragraphs we briefly outline the contributions herein.



This issue includes a paper from the group of T. Ichiye *et al.* [DOI: 10.1002/pro.2772] on the dynamics of the iron protein (FeP) from *Azobacter vinelandii* involved in nitrogen fixation. In this work a key question regarding the nature of the ionization state (+1/0) is examined using density functional theory, numerical Poisson Boltzman and elastic network analysis and the authors show that fluctuations, i.e., protein dynamics, might lead to variations in the reduction potential to be consistent with conflicting experimental observations.

Following on this theme of protein dynamics, a paper in this special issue by D.E. Shaw *et al.* [DOI: 10.1002/pro.2770] focuses on the question to what extent do a small number of native-like biases, in the form of native-residue pair contacts, provide acceleration of protein folding. By exploring folding of ubiquitin with a varying number of residue-residue contacts, work from this paper demonstrates that speed-ups of approximately ten-fold can be achieved in folding.

DOI: 10.1002/pro.2851
Published online in Wiley Online Library (wileyonlinelibrary.com).

Montelione *et al.* [DOI: 10.1002/pro.2774] contributes a review of the community resource of experimental data for NMR-X-ray structural pairs. The author provides an overview of the protein structure initiative and the goal of developing a database of paired protein structures from NMR and X-ray for comparative studies, such as that performed by R. Levy with G. Montelione, demonstrating the general agreement between models derived from both structural approaches.

In a contribution from the laboratory of E. Arnold *et al.* [DOI: 10.1002/pro.2776], the development of a modified DNA aptamer that binds to HIV-1 reverse transcriptase (RT) is described. A key outcome discussed in this work is the finding that the RT-aptamer complex is catalytically active and as such can serve as a platform to study fundamental RT mechanisms and to develop anti-HIV inhibitors through fragment screening and related methods.

The role of solvation and the balance of interactions between protein intra-molecular interactions and protein-solvent energies in the heat denaturation of the Trp-cage mini-protein was explored with explicit solvent molecular dynamics by N. Matubayasi *et al.* [DOI: 10.1002/pro.2754]. It was observed that the solvation free energy is anticorrelated with the protein-intramolecular energy, independent of temperature.

In a contribution from D. Zuckerman *et al.* [DOI: 10.1002/pro.2738] and co-workers, a method of estimating the first passage time distribution from weighted-ensemble simulations and non-Markovian analysis was described. The authors argue that a non-Markovian analysis approach with a sufficient subset of history information, can provide robust estimates of unbiased mean first passage times from weighted-ensemble simulations.

The feasibility and possible mechanism of substrate channeling between human dihydrofolate reductase and thymidylate synthase is explored from a theoretical perspective in a paper by N. Wang *et al.* [DOI: 10.1002/pro.2720]. The authors examines whether substrate channeling, as hypothesized to be in effect in protozoa for this enzyme pair, is possible for the human analogues using a combination of protein-protein docking, electrostatics calculations and Brownian dynamics simulations. The results of the modeling suggest that channeling with an efficiency close to that seen in protozoa is possible.

In work aimed at better assessing the quality of modern force fields for molecular modeling, D. Case *et al.* [DOI: 10.1002/pro.2713] and co-workers explore the structure and fluctuations of lysozyme in the triclinic lattice using molecular dynamics on the microsecond timescale. Comparing several current force fields the authors find all perform comparably, but the Amber ff14SB results yield an average structure that is slightly closer to the deposited X-ray model.

In a paper from M. Pettitt's *et al.* [DOI: 10.1002/pro.2749] laboratory, a theoretical analysis of the forces that underlie protein folding, collapse and aggregation are examined for glycine oligomers. Work presented in this paper suggests that hydrogen bonding interactions are not contributing to processes that tend to decrease the exposure of the polypeptide to solvent, but instead the backbone carbonyl dipole interactions are the key drivers of these processes.

Similar in spirit to the theme from the D.E. Shaw paper, J. Onuchic *et al.* [DOI: 10.1002/pro.2758] contributes a manuscript that explores the power of mean-field direct coupling analysis (mfDCA) in extracting coevolutionary residue-residue interaction information from the statistical analysis of large collections of protein sequences. Using inferred pairwise statistical coupling information from mfDCA to inform molecular dynamics simulations employing coarse-grained structure-based models the authors show that the folding mechanism of ribosomal protein S6 and several of its circular permutants as well as SH3 can be deduced in quantitative agreement with experiment, in general agreement with Shaw's observations that a few native interactions inferred from statistical analysis can greatly enhance for yield of natively folded structures in explicit solvent molecular dynamics.

The analysis of single molecule force spectroscopy experiments to extract timescales, barrier heights and diffusion constants is discussed in the paper by D. Makarov *et al.* [DOI: 10.1002/pro.2727]. In this paper the author provides a model that aims to guide researchers carrying out such experiments in their accounting for instrument parameters, and the importance of the relative molecular elasticity of the molecule being examined and that of the linker that attaches the molecule to the instrument probe. Marakov provides "optimal" instrumental parameters that enable determination of "instrument-free" molecular dynamics.

In work presented by R. Elber *et al.* [DOI: 10.1002/pro.2723], a theoretical framework for discovering or designing protein switches, i.e., protein sequences that switch between dramatically different structural motifs in response to small sequence changes, is described and examined for discovery of a putative switch pair between an all alpha and an alpha plus beta fold. A sequence pair is predicted and put forward as a challenge for experimental evaluation.

Integral equation theory, 3D-RISM/RISM, is used by T. Rungrotmongkol *et al.* [DOI: 10.1002/pro.2718] to examine the binding affinity of oselramivir, an anti-viral compound, to influenza B neuraminidase and three resistance mutants. This study shows the classic compensation between desolvation and receptor-ligand interactions in yielding the

ultimate binding affinity. Further, it is also found that the theoretical models are able to reproduce the trends in binding affinities between the drug and the neuraminidase variants.

By exploiting the growing structural coverage of proteomes, template-based approaches to predict regions on a protein surface likely to bind other proteins becomes feasible and such approaches are described in a paper from the laboratory of B. Honig *et al.* [DOI: 10.1002/pro.2744]. The new method, PredUs 2.0, utilizes a Bayesian approach combined with the template-based scoring in predecessor methodology PredUs used to analyze known binding sites in structurally similar proteins to predict interfacial residues. The findings indicate that PredUs 2.0 significantly outperforms PredUs and other published interface prediction methods.

In work from the C. Chang *et al.* [DOI: 10.1002/pro.2709] group, the importance of protonation states and proton transfer in pyridoxal 5′-phosphate (PLP)-chemistry in the tryptophan synthase (TRPS) enzyme family is explored using solid-state NMR informed molecular dynamics simulations. The study of interactions in the active site of TRPS shows that functional groups on the reacting substrate, such as the phosphoryl group, pyridine nitrogen, phenolic oxygen and carboxyl group, of each PLP-bound intermediate play a crucial role in constructing an appropriate molecular interface with TRPS, and the protonation states of the ionizable groups on the PLP cofactor modulate interactions between the enzyme and substrate.

The eukaryotic translation initiation factor 4E (eIF4E) interacting with eIF4E binding proteins (4E-BP) represents a template for the inhibition of eIF4E, relevant to the treatment cancer and autism disorders, including the Fragile X syndrome. In work described by E. Gallicchio* *et al.* [DOI: 10.1002/pro.2708] and coworkers, an atomically detailed model of the complex between eIF4E and a peptide fragment of a 4E-BP, the cytoplasmic Fragile X interacting protein (CYFIP1) is described. The model, from computer simulations with enhanced sampling utilizing a novel alchemical replica exchange approach, provides a basis for and suggests an alternative strategy for the design of eIF4E inhibitor peptides.

In a paper contributed by the W. van Gunsteren *et al.* [DOI: 10.1002/pro.2695] group computational methods to construct conformational free energy differences are examined in a comparative study. Using a test system for the interconversion of a hexa-β-peptide from a left-handed to a right-handed helix in methanol, the authors demonstrate that one needs to take care in converging umbrella sampling simulations, since poor convergence and overlap of the resulting distributions in conformational space can yield imprecise estimates of the free energy difference. The work contrasts these approaches with direct counting approaches, such as employed in the enveloping distribution sampling approach, and suggest the later to be preferable from an ease of convergence standpoint.

C. Wong *et al.* [DOI: 10.1002/pro.2716] presents a paper examining the utility of fast simulations of conformational transition paths for elucidation of enzyme mechanism and drug discovery for protein kinases. Using transition path methods in the MOIL software package, conformational transitions between pairs of protein kinase structures are explored with the major finding being that conformations identified along the transition paths resemble experimental structures believed to form during enzymatic catalysis or to mimic drug bound conformations. These observations lead the author to suggest that such simulations may be useful in gaining initial insights into the enzymatic mechanisms, pathways of conformational transitions of proteins kinases, or structures for structure-based drug design.

In work presented by the group of C. Post *et al.* [DOI: 10.1002/pro.2753], three implicit solvent models, GBMV II, FACTS, and SCPISM, were evaluated for their ability to mimic explicit solvent. By comparing the conformation ensembles, dynamics and electrostatic interactions of the Src SH2 domain and the Lyn kinase domain from molecular dynamics simulations utilizing the three different implicit solvent models, the authors find the folded proteins expand. Also, using several conformational and structural fluctuation based metrics they observe that, compared to explicit solvent simulations with TIP3P solvent, both GBMVII and FACTS show that the Src SH2 domains sample similar global conformations, and show similar ion-pair distance distributions for solvent-exposed side chains. Similar correspondence is not seen for the SCPISM model. For the non-globular Lyn kinase domain, with a bi-lobal structure, none of the implicit solvent models maintain the fidelity of partially exposed ion-pair interactions and lobe structure, suggesting some caution in the blind application of these approaches to non-globular proteins.

In a paper from B. Roux's *et al.* [DOI: 10.1002/pro.2731] group the underlying mechanism of mutation-induced pathologies in kinase activity was explored for the Src tyrosine kinase, where the mutation of a highly conserved tryptophan to alanine (W260A) is correlated with increased activity of the kinase, linking it to a role in cancer cell growth. From a combination of atomic scale simulation and umbrella sampling together with a simulation informed kinetic model,

*The article by Gallicchio et al. was published in an earlier issue of Protein Science [Di Marino D, D'Annessa I, Tancredi H, Bagni C, Gallicchio E (2015) A unique binding mode of the eukaryotic translation initiation factor 4E for guiding the design of novel peptide inhibitors. Protein Sci 24:1370-1382].

these authors suggest that small differences in the populations of the active-like state are sufficient to yield the accelerated trans-autophosphorylation activity observed experimentally.

Olson *et al.* [DOI: 10.1002/pro.2733] and his coworkers describe in this issue the implementation of a covalent docking scheme in Autodock4. They describe two different approaches for setting up docking studies for covalently bound ligands and note that of the two, a two-point attractor or flexible side chain scheme, the flexible side chain performed better in docking of a set of protein-ligand complexes.

In a paper from the B. Brooks *et al.* [DOI: 10.1002/pro.2755] group the authors combine their constant pH frameworks with the Bennett acceptance ratio method to explore pH dependent ligand binding. They suggest, based on findings from their study that their approach enables some of the first combined pH and ligand binding calculations. They also compare various means of treating long-range electrostatic interactions in constant pH simulations.

Szabo *et al.* [DOI: 10.1002/pro.2722] and coworkers present a study of kinetic processes for multisite reactive systems using numerical solutions to reaction-diffusion equations. They show that when an enzyme modifies multiple sites on a substrate, the influence of the relative diffusive motion of the reactants cannot be described by simply altering the rate constants in the rate equations of chemical kinetics. Instead, the corresponding new rate constants depend on capture and rebinding probabilities that an enzyme-substrate pair reacts at a neighboring site diffusing away. Additionally, they show that similar results arise from a phenomenological approach based on the introduction of transient encounter complexes into the standard kinetic scheme and then eliminating them using the steady-state approximation.

The protein structure prediction potential ASWEM from the group of P. Wolynes *et al.* [DOI: 10.1002/pro.2751] is augmented with electrostatic interactions to explore the role of these long range but generally non-specific forces in protein folding and binding. The authors suggest, based on folding and binding simulations using this new model, that electrostatic interactions may sometimes play a facilitating role in folding/binding processes but in other instances can lead to frustration. These points are illustrated with studies of the KIX-pKID complex and the DNA-binding protein FIS.

Exploiting the novel computational approach of generalized orthogonal space sampling (gOST), the research group of W. Yang *et al.* [DOI: 10.1002/pro.2789] explored the free energy landscape of ionic salt bridge formation in a model Asp-Arg dipeptide. From these calculations it was concluded that the opposing forces of solvation and ionic pairing of oppositely charged side chains may be easily modulated by local influences that alter the solvent structure and thereby facilitate salt bridge formation.

Employing a hybrid QM/MM approach, R. Freisner *et al.* [DOI: 10.1002/pro.2819] and his colleagues present a paper that examines the role of dispersion and localized orbital corrections in quantitatively capturing the transition-state barriers for enzymatic processes utilizing transition metals. They find that for two important reactions, the hydroxylation of camphor by cytochrome P450 and C-H bond activation in methane monooxygenase that an appropriate combination of these corrections allows the calculated barriers to come into accord with experimental estimates to within 1.1 kcal/mol.

Experimental studies on fibril formation and conformational preferences in β-synuclein (βS) are reported from the Baum lab *et al.* [DOI: 10.1002/pro.2798] in this issue. Using NMR residual dipolar coupling and secondary structure propensities deduced from chemical shifts together with a number of biophysical techniques, Baum and co-workers demonstrate that the missense mutation of βS, P123H, leads to a more flexible C-terminal tail and subsequent increased facility to form fibrils upon coincubation with α-synuclein, whereas wild-type βS does not form fibrils. Thus, this mutation, identified in patients who develop Dementia with Lewy Body disease, and its link to conformational flexibility in the C-terminus may provide a basis for therapeutic intervention in Parkinson's disease.

Carol B. Post
Department of Medicinal Chemistry and
Pharmacology Purdue University West
Lafayette, IN 47907
Phone: (765)494-5980
FAX: (765)494-1414

Charles L. Brooks III
Department of Chemistry and Biophysics
University of Michigan
930 North University Avenue, Chem 2006
Ann Arbor, MI 48109
Phone: (734)647-6682
FAX: (734)647-1604