Featured DPMB members

Stefan Wallin, Ph.D.

Assistant Professor
Department of Physics and Physical Oceanography, Memorial University
swallin@mun.ca
https://www.physics.mun.ca/~stefan/

Biography

Stefan Wallin received his PhD degree in Theoretical Physics in 2003 from Lund University. Thereafter he held a postdoc position at the University of Toronto (2003-2005) and another at Harvard (2006-2008). From 2008-2015, he was an Assistant Professor at Lund University. He joined Memorial University of Newfoundland in 2015.

Simulating proteins with coarse-grained models and alternative sampling methods: folding, fold switching and conformational disorder

My research focuses on the development of methods for biomolecular simulations and their application to biologically relevant systems. To this end, I use techniques from various fields of physics, especially computational and statistical physics. In particular, my focus is on the folding, interaction and evolution of proteins. While much progress has been made towards an understanding of protein folding through advances in both experimental and computational techniques, there are fundamental questions that remain unresolved (see e.g. Dill and MacCallum, Science 338 1042, 2012).

Since joining Memorial University in 2015, I have worked on one of these unresolved questions, namely how variations in the amino acid sequence impact the folding process. I have also investigated the biophysical properties of so-called “metamorphic” proteins, which are proteins that exhibit a remarkable ability to reversibly switch between alternative native folds. In terms of method development, I have worked on the problem of how to interpret low-resolution electron density maps from cryogenic electron microscopy. The method we developed was used to characterize the structure of high-density lipoproteins.

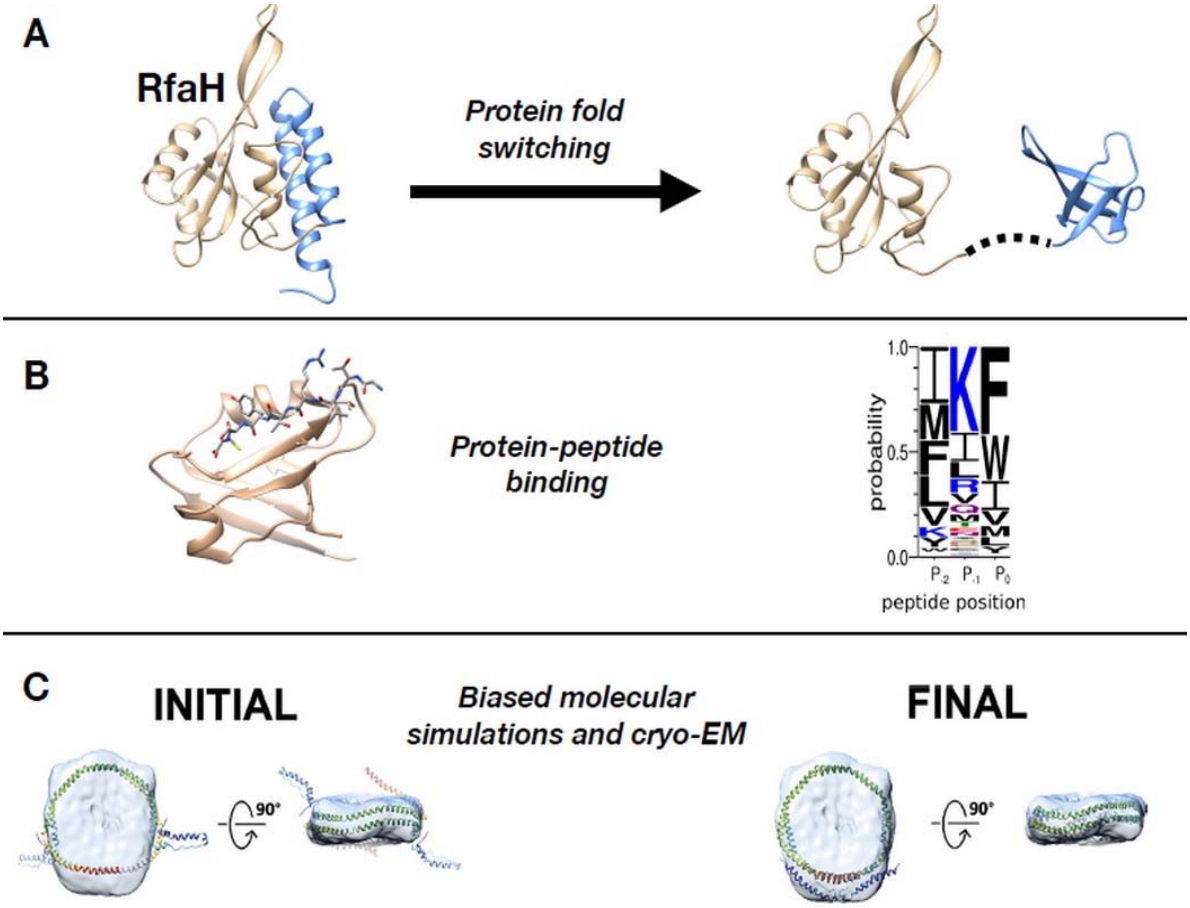
Proteins are generally challenging targets for molecular simulation methods because of their relatively large size, conformationally specific behavior and range of timescales involved. Typical folding times of small, single domain proteins alone span microseconds to seconds. To make things more difficult (as well as more exciting), current biophysical experiments on proteins increasingly emphasize the role of large-scale conformational transitions for how functions are carried out, such as the coupled folding-binding processes of intrinsically disordered proteins.

A substantial part of my research is therefore devoted to the computational challenges of biomolecular simulations. I tackle these in two ways. On the one hand, by developing models at various levels of resolution, acknowledging that coarse-grained models sometimes have limited transferability across different systems due to the diverse behaviors of proteins. On the other hand, by developing alternative methods for conformational sampling. I focus in particular on advanced Monte Carlo-based techniques, which can speed up the calculation of equilibrium properties of biological molecules and systems, such as folding free energy surfaces, by circumventing the inherent time-dependence of traditional molecular dynamics simulations. Recently, I have developed a generalized-ensemble method for efficient exploration of the equilibrium behavior of

multiple amino acid sequences, which can be used to study sequence effects in either binding or folding.

What are "metamorphic" proteins?

An interesting new development in the protein folding field has been the discovery of an increasing number of natural and engineered proteins that are able to reversibly switch between different three-dimensional structures. Because these transitions involve major changes in secondary structure content, hydrophobic core packing and overall shape, these proteins have been termed "metamorphic" (Murzin A G, Science 320, 1725-1726 2008). Their behavior is entirely in contrast to the traditional view of proteins, which holds that proteins have a unique native structure. Fold switching is typically triggered by an external signal such as ligand binding, a change of pH or – as in evolution – by mutational changes. We are interested in understanding the biophysical properties of these large-scale conformational transitions: What factors control the (delicate) free energy balance between different folded states? Why is fold switch so abrupt? How can different types of mutational changes drive proteins to switch folds? An example of a partial but dramatic fold switch is given by the Escherichia coli protein RfaH (see figure). The C terminal domain of this protein transitions from an α-helix bundle to a β-barrel, triggered by the separation of this domain from the rest of the protein. A graduate student in my group, Adekunle Aina, investigated conformational and stability properties of this protein. From simulations of the various



Three current areas of research. *(A) The C-terminal domain (blue) of the protein RfaH undergoes a spontaneous and reversible fold switch from an α-helical bundle to a β-barrel. (B) The peptide binding specificity of the protein GRIP1 PDZ7 determined using our computational peptide screening algorithm. (C) A model 3D-structure of discoidal high-density lipoprotein (HDL) obtained from low-resolution cryo-EM data and biased molecular simulations.*

structural forms of RfaH, including fragments and the full-length protein, Adekunle found a markedly reduced stability of the fold-switching region of RfaH. In particular, this suggests a way to search for as-of-yet undiscovered metamorphic proteins based on structural information. In his MSc project, he also applied a coarse-grained model to systematically study the folding properties of sequences in the space "in between" different protein folds. In order to make this computationally possible, he developed and studied a novel simulation algorithm that allows the equilibrium behavior of multiple sequences to be determined in a single simulation trajectory.

Disordered protein and peptide-protein interactions

Another major interest in my group is so-called intrinsically disordered proteins (IDPs). These proteins are special in that they persist as a dynamic ensemble of interconverting conformations even under native conditions, and yet they are fully functional. Somewhat counterintuitively, IDPs are often involved in molecular recognition, i.e., making specific interactions with other biomolecules. In these situations, IDPs typically undergo a folding transition as they bind their target molecule. We have found, however, along with other computational groups, that IDPs can exhibit substantial structural heterogeneity even in their final bound complexes. This behavior has major implications for how IDPs realize specific interactions. It also has practical implications for binding free energies calculations, as conformational entropy can be an important factor to binding.

We have also developed an algorithm that characterizes the specificity of a given peptide-binding pocket through molecular simulations. The algorithm works such that sequences s can be generated according to a Boltzmann weight, i.e., $P(s) \propto \exp(-\Delta F(s)/RT)$ where ΔF is the binding free energy at temperature T. This way strongly binding peptide sequences are frequently generated while poor binders are supressed. Importantly, this allows the specificity of peptide-binding sites to be explored even as conformational entropy factors are taken into account. We are currently exploring the possibility of adapting our algorithm for the discovery of unknown peptide-binding sites on protein structures.

A computational method for interpreting data from cryo-electron microscopy

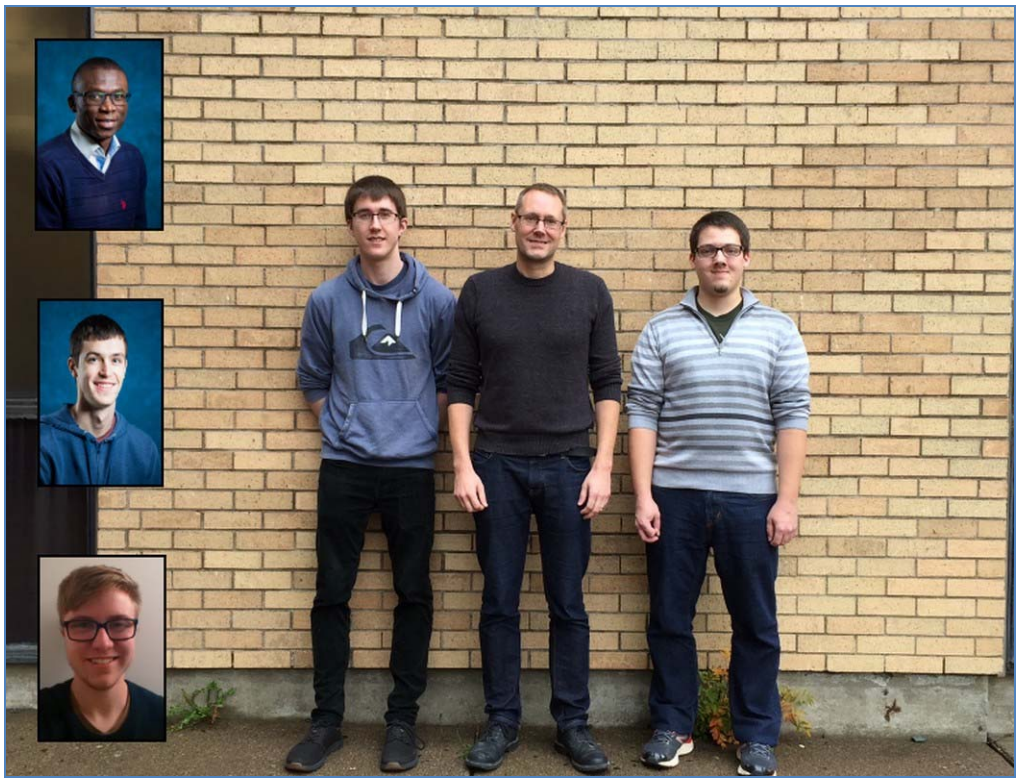
In a collaborative project with Karolinska Institute (Caroline Jegerschöld), and Lund University (Jens Lagerstedt), we developed a computational method for structural interpretation of data from cryogenic electron microscopy (cryo-EM). The determination of biomolecular structures has for a long time been dominated by X-ray crystallography and nuclear magnetic resonance spectroscopy. However, single-molecule cryo-EM is emerging as a rivalling technique (see Ewen Callaway Nature 525 2015). In this technique, biomolecular assemblies are flash-frozen in a thin film of vitreous ice and imaged at different angles. Many low-resolution images are combined into one or a few three-dimensional electron density maps, which must be interpreted through molecular modeling.

In this project, an undergraduate student in my group, Peter Gysberg, developed a method that combines our all-atom Monte Carlo framework for protein simulations with an auxiliary energy term that scores the correlation between experimental and calculated electron densities, thereby steering simulations toward structures that fit the underlying data. We applied this method to cryo-EM data on high-density lipoproteins (HDL), obtained from our collaborators. HDL is a blood plasma particle involved in lipid and cholesterol transport. Because its abundance in the blood is inversely correlated with the risk of cardiovascular disease, HDL is sometimes referred to as the "good" cholesterol. The large size and high conformational flexibility of HDL makes it challenging for structural studies and its structure is debated. By combining our simulation results with other data from in the literature, we recently reported new structural models for this particle (see figure).

Key publications:

1. Aina A, and Wallin S. Multisequence algorithm for coarse-grained biomolecular simulations: exploring the sequence-structure relationship of proteins. Journal of Chemical Physics 147 095102 (2017).
2. Staneva I, Huang Y, Liu Z, and Wallin S. Binding of two intrinsically disordered peptides to a multi-specific protein: A combined Monte Carlo and Molecular Dynamics study. PLoS Computational Biololy 8, e1002682 (2012).
3. Bhattacherjee A, and Wallin S. Exploring protein-peptide binding specificity through computational peptide screening. PLOS Computational Biology 9, e1003277 (2013).
4. Zhu L, Petrlova J, Gysbers P, Hebert H, Wallin S, Jegerschöld, C J and Lagerstedt J O. Structures of Apolipoprotein A-I in High Density Lipoprotein generated by Electron Microscopy and Biased Simulations. Biochem Biophys Acta — General Subjects 1861 2726-2738 (2017).

My group at Memorial University



From left to right: Nicholas Robichaud (summer, honors projects), Stefan Wallin, Aidan Tremblett (MSc). Insets from top to bottom (previous members): Adekunle Aina (MSc; now at UBC), Peter Gysbers (summer project; now at UBC), Daniel Trotter (summer, honours projects; now at UOttawa). Not shown: Ryan Wilkins (summer, honours projects; now at UOGuelph).