Homework 0

Bio

I am a first-year graduate student in the Department of Bioengineering. When I began college a few years ago, I was interested only in studying biology. Over the past few years, though, I have learned that I am also very interested in programming, computer science, and math. My goal in my research is to combine these fields with biology to make biological research more efficient.

Specifically, my focus is in systems biology. I am very interested in the computational modeling of complex biological systems using theoretical and data-driven models. Some areas of biological research that are currently investigated using computational methods, like protein folding, require extremely complex molecular dynamics simulations, and are often run using parallel computing. As a result, I hope to learn a lot about the various uses of parallel computing, design of parallel programs, and more biological applications of parallel computing in this class. I hope that by the end of this course I will be able to think of new ways to use parallel computing in biological research.

Long-timescale Models of Membrane Protein Translocation

In Zhang and Miller 2012, the authors use parallel computing to run coarse-grained models showing how proteins integrate into the Sec translocon in the membrane. For simplicity, the authors focus on only single-spanning proteins to remove the complexity of multiple integrations. There are two types of integration possible: type II, or cytoplasmic N-terminus, and type III, cytoplasmic C-terminus. In their models, the authors include three methods of integration, two for type II (loop and flipping) and one for type III.

The major advancement in this paper is that the authors have created models capable of showing integration over minutes-long timescales, which is at least an order of magnitude improvement over models of the same type (cite). There are a few major assumptions made in this model to decrease the computational power needed to model the protein integration:

- Primarily, the authors model only proteins which enter normal to the Sec translocon and membrane, which greatly reduces complexity
- Coarse-grained models prevent the resolution of interactions of specific amino acid residues with the Sec translocon
- Other membrane-bound proteins are not considered

Overall, the authors present this work as a very minimalist interpretation of protein integration in the translocon; the main importance of this model is that it allows researchers to model relevant timescales (minutes) for membrane protein integration. Simple models often have great utility. However, the authors did not present a comparison of their model’s results to physical data. As a result, it is extraordinarily difficult to understand the accuracy or practical utility of this model.

Little information is given about the specific conditions on which this model was run. The authors do note that the CPU time needed to calculate each protein’s trajectory ranges from 0.2-
10 hrs, though the specifications of the system used and whether this is while running in parallel are not discussed. Further information on this topic would also be welcomed.