High-Accuracy Stochastic Methods for Breakthrough Electronic Structure Calculations

Key Idea and NCSA Skills/Resources/Interests:

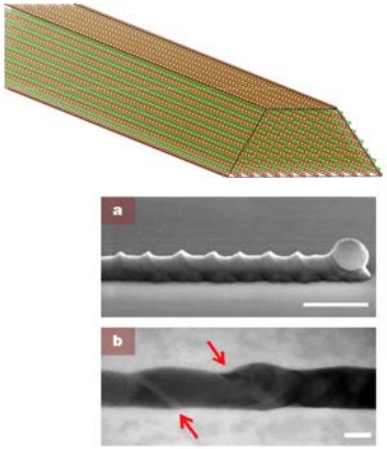
- Quantum Monte Carlo (QMC) methods are benchmark accuracy stochastic methods for electronic structure, and demonstrate near-linear parallelization to millions of cores. They are well-established in the physics community.
- We are interested in extending the QMC method from a physics tool to a widely applied tool for understanding real engineering materials (*e.g*. nanowires with twin plane defects, complex semiconductors, *etc.*). Current state of the art focuses on bespoke calculations, which limit systematic application to engineering problems.
- **Needed NCSA skills center on data management**: large volumes of computations will be key to overcoming the learning curve and establishing the framework for application of a new method
- We are interested in developing and maintaining a **live database of QMC results** for use by the entire international simulation community. This living data base will be a **key enabler** for developing the protocols and frameworks for the wide scale application of this method. We intend that this one-year work will form the basis for a large Illinois/NCSA proposal, which we anticipate to be of substantial interest to the US DOE.

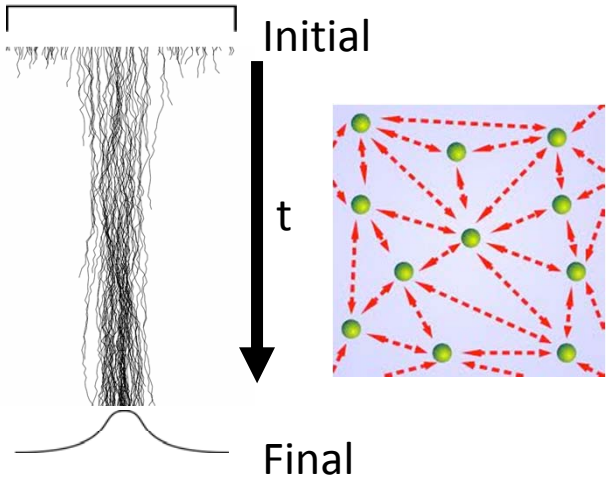
Illinois Faculty Team:

Elif Ertekin, MechSE

Lucas Wagner, Physics

Xiuling Li, ECE





contact: ertekin@illinois.edu

Andrew L Ferguson

Assistant Professor of Materials Science and Engineering
Affiliated Assistant Professor of Chemical and Biomolecular Engineering
Affiliated Assistant Professor of Computational Science and Engineering

204 Materials Science and Engineering Building (MC-246)
University of Illinois at Urbana-Champaign
1304 W. Green St., Urbana, IL 61801

alf@illinois.edu
http://ferguson.matse.illinois.edu

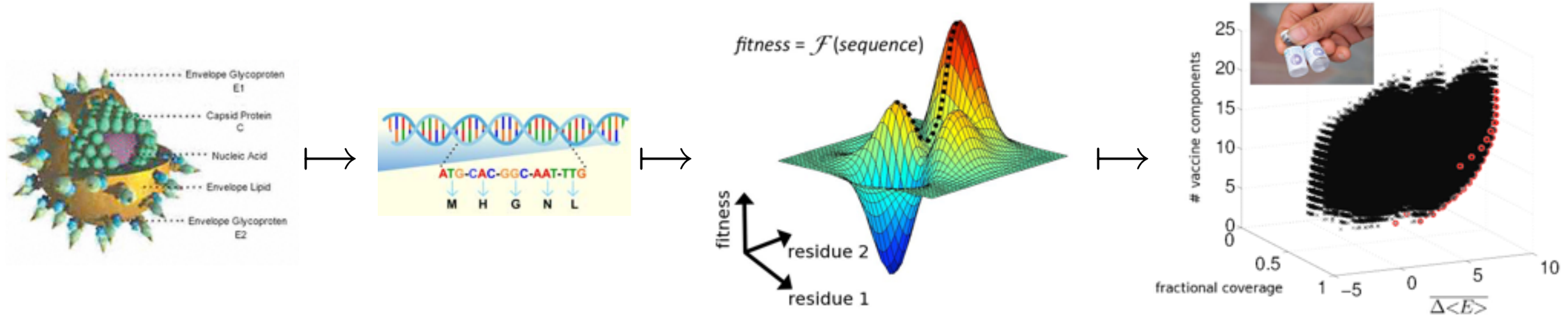
Computational Design of Hepatitis C Vaccines

Prof. Andrew Ferguson, MatSE

NCSA Thematic Area: Biological & Health Sciences

PROJECT SYNOPSIS

- Bayesian inference of hepatitis C virus fitness landscapes from clinical sequence databases
- Fitness landscape prescribes viral replicative capacity as a function of proteome amino acid sequence
- Landscape described by a Potts spin glass Hamiltonian
- Model parameters fitted by iterative Monte Carlo fitting of model predictions to clinical data
- Quantitative landscapes reveal viral “soft spots” and guide rational vaccine design — no vaccine is yet available



PROJECT NEEDS

- Computational bottlenecks in Monte Carlo sampling limits us to single viral proteins
- Full proteome landscapes require (i) large-scale code parallelism and (ii) supercomputing infrastructure
- Codes are CPU and GPU parallelized but inexpertly and inefficiently — professional support invaluable
- Extension to full hepatitis C proteome — NCSA computing resources vital
- Success will massively accelerate discovery of viable vaccine candidates, alleviating the suffering of 170 million infected persons worldwide = 3% of global population

Constrained Optimization of Agent-based Disease Models

- I have the conceptual framework and parameters for agent-based/hybrid disease models in livestock herds
  - http://www.aaai.org/ocs/index.php/IJCAI/IJCAI11/paper/view/3304
- I want to work on designing these models to allow constrained optimization, ideally over multiple state spaces
  - *Allow farmers to prioritize disease-control spending*
- I need expertise in building agent-based models and interest in developing novel optimization techniques

Rebecca Smith, DVM MS PhD
Assistant Professor of Epidemiology
Department of Pathobiology
College of Veterinary Medicine
rlsdvm@Illinois.edu
217-300-1428