



## Science Highlight

Using Beagle in an undergraduate biology course

*Esmael J. Haddadian*

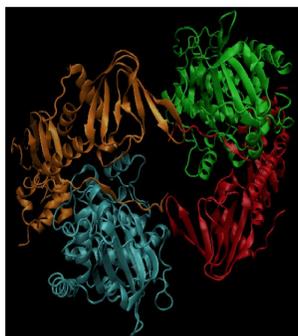
In addition to research, a core component of Beagle's mission is education, as a hands-on platform for introducing students in the biological sciences to quantitative methods and high performance computing. This past academic year, Beagle (in combination with the Research Computing Center's cluster, "Midway") was used in two undergraduate courses at the University of Chicago. Both of these two courses have attracted a large population of our students that have grown in numbers over years.

**Multiscale Modeling of Biological Systems II (BIOS 106030):** This course is the second part of the two-quarter long course that aims to introduce students majoring in quantitative sciences to elements of modern biology via quantitative modeling. It covers topics ranging from water solvation chemistry and amino acids, to the structure and function of proteins, DNA, and RNA. The students learned how to visualize macromolecules and measure their basic properties and to model their physical movements by means of molecular dynamic simulations running at university's super computers.

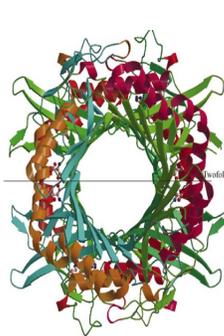
Students in this course learned about high performance computing (HPC) via two lectures comprising an introduction to HPC and practical Linux by Lorenzo Pesce and Joe Urbanski from the Beagle team. Students also learned to use VMD (<http://www.ks.uiuc.edu/Research/vmd/>), a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3D graphics and built-in scripting, and NAMD (<http://www.ks.uiuc.edu/Research/namd/>), a parallel molecular dynamics code designed for high-performance molecular dynamics simulation of large biomolecular systems.

For their final project the students chose a protein system to simulate and analyze using NAMD and VMD packages. Students used a total of 360,000 core hours on their protein systems, and used Crerar Library's Kathleen Zar room (part of the Research Computing Center's facilities) to view the simulations in 3D. It was very pleasing to see the student fascinations when they saw the movements of their proteins at the atomistic detail. During finals week, students presented their results to the class and then submitted a write-up.

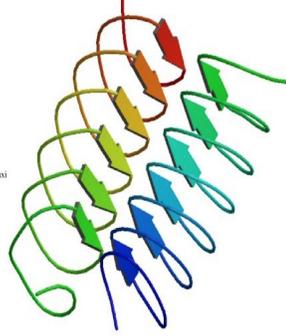
### Examples of final projects were:



Protein Kinase A



Urate Oxidase



Beetle anti-freeze Protein

### Resources:

#### Beagle Wiki

Get detailed usage information from the Beagle team

#### Beagle Support

Contact the Beagle experts for help

#### Globus

Get started moving files to/from Beagle using this fast service

#### Other CI resources

Learn about other computing resources available at the Computation Institute



### Training:

#### Intro to Beagle

September 14th, 10AM  
Room 240A  
at the Computation Institute of the University of Chicago

Topics will include:

- General overview of Beagle2's team and the Computation Institute
- Overview of Beagle2's Cray XE6 system architecture
- Basic access and navigation operations
- Using compilers and applications
- Appropriate use of local and network filesystems
- Submitting jobs and monitoring jobs
- Data transfer
- Specific topics requested by users (e.g., using R or Matlab on Beagle2)

Beagle2 Events To learn more about Beagle2 trainings.

**An Introduction to Bioinformatics and Proteomics (BIOS 26120):** This is a one-quarter course devoted to understanding biological data and the models and computational techniques used to make sense of them. The course covered biological databases, algorithms for sequence alignment and phylogenetic tree building, and large-scale protein analysis techniques. Students learned how to use our in-house supercomputer facilities to process and analyze next generation sequencing (NGS) data.

Similarly, to the previous class, students received introductory HPC training including its application to the field of bioinformatics, how to practically use Beagle and its operating system. For their final project, students downloaded 1000 genomes raw NGS files, aligned raw NGS reads to the reference human genome, identified single nucleotide variant (SNV) and insertion-deletion (indel) mutations, and annotated these variants with biologically pertinent information. They then explored the annotated variants for their potential roles in human diseases.

Overall students enjoyed using Beagle and Midway super computers as part of their coursework. "The HPC part of the class is useful because computation is a tool used by real researchers to make the discoveries studied in the rest of the class," one student said. "It's as important as discussion of lab methods is to a Chemistry class."

"The HPC aspect was the best part of the class," another student said. "While learning biological facts was interesting, it was through running the simulations that I feel I actually got some understanding of how computational biology is done, and it was fun to conduct a simulation on a protein of your own choosing. It wouldn't be possible if we didn't have access to a supercomputer large enough to allow each of us to run around 40 hours of simulation."

I am planning to use Beagle again in the coming academic year for the same classes. I am hoping to expand the educational role of Beagle with the support of Lorenzo and Joe.

## Beagle2 Related Publications

Z. Lou, J. Reinitz

*Parallel simulated annealing using an adaptive resampling interval.*  
Parallel Computing; 53: 23-31

R.C. Cockrell, M. E. Stack, G. An

*Supercomputing ulcerative colitis-associated cancer simulations to bridge mechanism with epidemiology*  
Digestive Disease Week, San Diego, CA, 5/21/2016

R.C. Cockrell, G. An

*Characterizing the behavioral landscape of sepsis: supercomputing simulation of 40 million in silico patients*  
Society of Critical Care Medicine Annual Congress, Orlando, FL 02/03/2016

R.C. Cockrell, G. An

*Supercomputing sepsis simulations for in silico outcome prediction*  
Academic Surgical Congress 2016, Jacksonville, FL 02/03/2016

R.C. Cockrell, G. An

*Investigating the development of ulcerative colitis-associated cancers through an agent-based model*  
Academic Surgical Congress 2016, Jacksonville, FL 02/03/2016

L. Waldron, J.D. Steimle, T.M. Greco et al.

*The Cardiac TBX5 Interactome Reveals a Chromatin Remodeling Network Essential for Cardiac Septation*  
Developmental Cell 2016; 36(3), p262-275, 8 February 2016

Searle Chemistry Laboratory  
5735 South Ellis Avenue  
Chicago, IL 60637



THE UNIVERSITY OF  
CHICAGO  
MEDICINE &  
BIOLOGICAL  
SCIENCES



Argonne  
NATIONAL LABORATORY



THE UNIVERSITY OF  
CHICAGO

Help E-Mail:

[beagle-support@ci.uchicago.edu](mailto:beagle-support@ci.uchicago.edu)

Beagle2 CI Website:

[beagle.ci.uchicago.edu](http://beagle.ci.uchicago.edu)