Docking@Home

Delaware, February 1, 2009

Content

Information

Project name: Docking@Home (D@H)

Project Webpage: docking.cis.udel.edu

Project location: University of Delaware

Number of volunteers: 6,000

Number of active computers: 9,170

Welcome message from Michela Taufer

Dear Friends,

The past year has been a very busy year for the D@H team. The Docking@Home team moved to the University of Delaware in fall 2008 and was challenged with the purchase of new servers for the project and the installation of the software parts. Fortunately new members have been joining the original team composed by me, Trilce Estrada, and Abel Licon. Trilce's effort in the past year has been terrific and has made possible the upgrade of D@H from alpha to beta in such a short time. She is the main coordinator of the job submissions as well as the collection and analysis of returned data. We will share Trilce's results with you soon! Abel is still working in the group but his Master's thesis targets a different application than docking. We would have never been able to have CHARMM working so smoothly with BOINC without Abel's help in compiling and debugging the code.

Kevin Kreiser, a graduate student, and Robert Keller, an undergraduate student, joined the group in fall 2007 and worked on the D@H screensaver. Robert graduated in May 2008 and now is working for a local company. Kevin is now working on his PhD thesis and developing more intuitive human-computer interfaces for D@H so that our volunteers can better learn the science in the project and more easily interact with the project interface. Jason Parrot graduated at the beginning of 2008 and left Delaware. Before leaving, Jason helped us to put together and install the servers for the D@H project. Now Andy Roosen, the system administrator at our department, is helping us with the servers. Of course the science has always been an important component of our project. Obaidur Rahaman (Ramie) joined our team from the Department of Chemistry in spring 2008 and is helping Trilce and me to get real science out of our simulations.

“Since the project went public, on September 2008, 6,000 volunteers have joined D@H”

Brian Flad and the D@H webpage developer team supervised by Dr. Dannagal Young (University of Delaware’s Interactive Media) have impressed everyone with their terrific job on the webpage that was recently released. Brian is still working with us to make sure that our webpage is a valuable tool to share the science and the computer science in D@H with you all. A new member of the team has just joined the team, Philip Saponaro. Philip is a junior undergraduate student and is working with Trilce and Kevin to monitor the system and increase the graphical content of the webpage.

Of course, we are still collaborating with our partners Charlie Brooks and Roger Armen on the scientific part of the project. They moved to the University of Michigan last year. We have been having long conversations with Roger about the work done with D@H and slowly we are getting our results to a point that we can soon present them to you. David Anderson visited us in August 2008 and we are working with him on an emulator of volunteer computing projects that can help administrators to tune volunteer projects and make sure that volunteers get computation and credits.

Since the project went public, on September 2008, 6,000 volunteers have joined. We have sampled docking results for two of our three proteins with a simple docking algorithm. We are ready for the third protein and more accurate docking algorithms, so a lot of computation is available for our volunteers in the next 12 months. So far we have been collecting more than 10,000 jobs per day! This is terrific and only possible because of your valuable support, patience, and commitment. Many kudos to all our D@H volunteers.

Enjoy the rest of this newsletter!

Michela

Our protein-ligand complexes

The design of new pharmaceutical drugs relies on finding small molecules, called ligands, which dock into proteins and play an essential role in turning protein functions on or off. Finding candidate ligand structures in a wet lab is very expensive and time consuming. Computer simulations are used to accelerate this process and to reduce costs. However, the computational search for putative drugs (i.e., ligands that dock well in a protein) is a search in a large space of potential conformations and, therefore, is a very time- and compute-intensive process.

Docking@Home computationally searches the large space of potential ligand conformations, reducing the time and cost required to design new drugs by several orders of magnitude. In this search, the protein and ligand can have different size and flexibility and their characteristics can have an important impact on the computational cost. Docking@Home considered an initial pool of proteins representing three levels of flexibility. The proteins are: Trypsin (26 complexes), HIV Protease (30 complexes), and p38 alpha (12 complexes). Figure 1 graphically shows the three proteins and their relative flexibility with respect to each other.

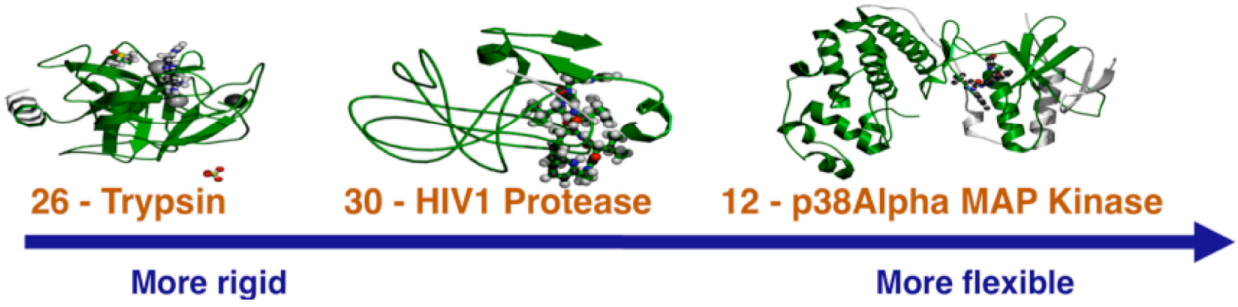


Figure 1: the D@H proteins

Trypsin

Trypsin is a protease that breaks down other proteins in the digestive system. Recent studies suggest that inhibitors of trypsin can have potential application in breast cancer treatment. In this terrible disease, trypsin-like proteases activate Protease-Activated Receptor-2 (PAR2), a protein in the tumor cell membrane. While activated, PAR2 causes the degradation of extracellular matrix (ECM) resulting in the spread of the tumor cell from one place to another (metastasis). Drugs can act as inhibitors by de-activating the trypsin-like protease and are therefore potential agents capable of stopping the spread of breast cancer [1]. Figure 2 shows the Trypsin and a small set of ligands that D@H considers in the simulations

[1] J. McIlroy, B. Cullen, G. Kay, J. Nelson, W. Odlings, R. Spence, and B Walker. A Novel Trypsin-like Enzyme in Breast Cancer. *Biochem. Soc. Trans.,* 22(1):19S, 1994.

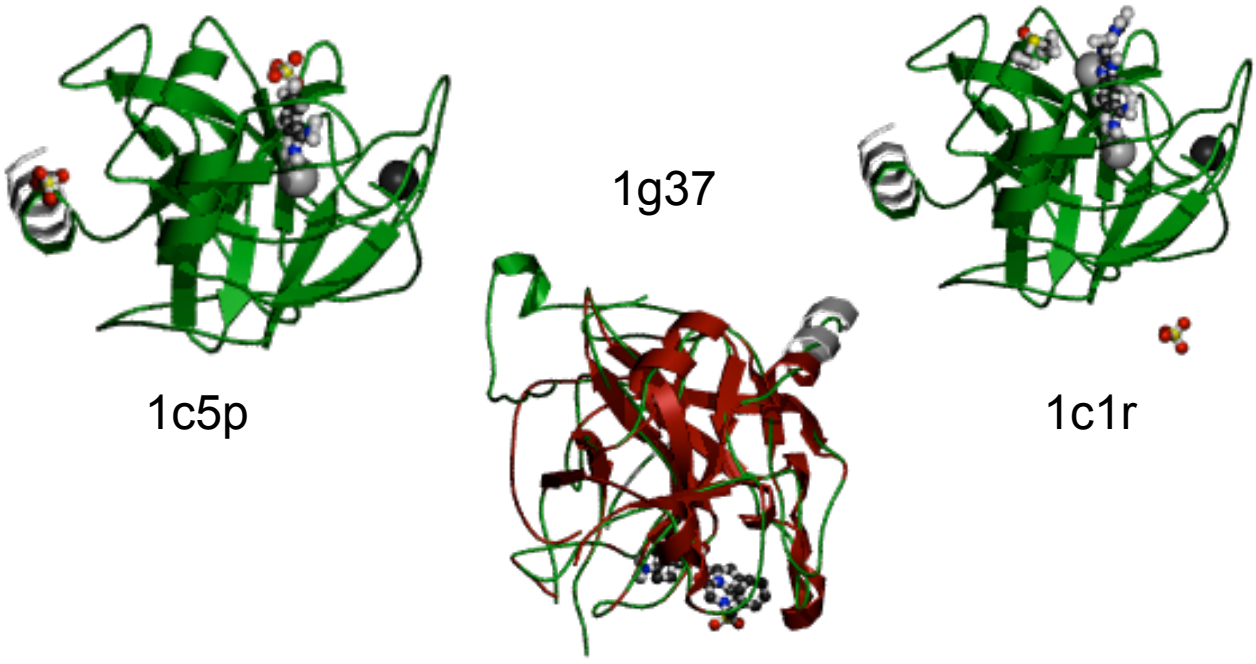


Figure 2: Three Trypsin complexes with their ligand docked in the protein

HIV Protease

HIV protease (HIV PR) is a protein in the HIV virus that is essential for its replication in human cells. During the process of building a new HIV virus inside the human cell, HIV PR cleaves some newly synthesized viral protein in a particular fashion. The cleaved pieces are required to build a mature HIV virus. HIV PR has been considered as a therapeutic target to prevent HIV proliferation. A drug that can bind to the active site of HIV PR or deactivate it in some way will eventually stop the replication of HIV virus in human cells. These drugs are called protease inhibitors. Several protease inhibitors like saquinavir, ritonavir, indinavir, and nelfinavir are available for the treatment of HIV infection [2]. Figure 3 shows the HIV protease and a small set of ligands that D@H considers in the simulations.

[2] L. Montagnier. Historical Essay: A History of HIV Discovery. *Science*, 298(5599): 1727 – 1728, 29 November 2002.

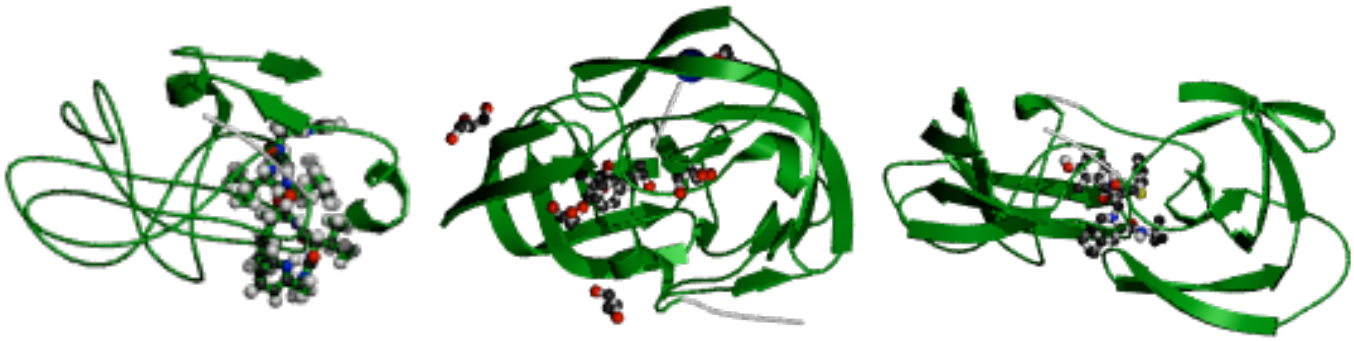


Figure 3: Three HIV protease complexes with their ligand docked in the protein

P38 alpha

P38alpha is also known as SAPK2a and MAPK14. It is involved in the regulation of cellular stress responses as well as the control of proliferation and survival of many cell types. Several promising compounds that inhibit p38 alpha are being investigated as potential therapies for arthritic and inflammatory diseases [3]. Figure 4 shows the p38alpha and a small set of ligands that D@H considers in the simulations.

[3] Lee, M. R.; Dominguez, C. MAP kinase p38 inhibitors: clinical results and an intimate look at their interactions with p38alpha protein. *Curr Med Chem* **2005,** 12, 2979- 94.

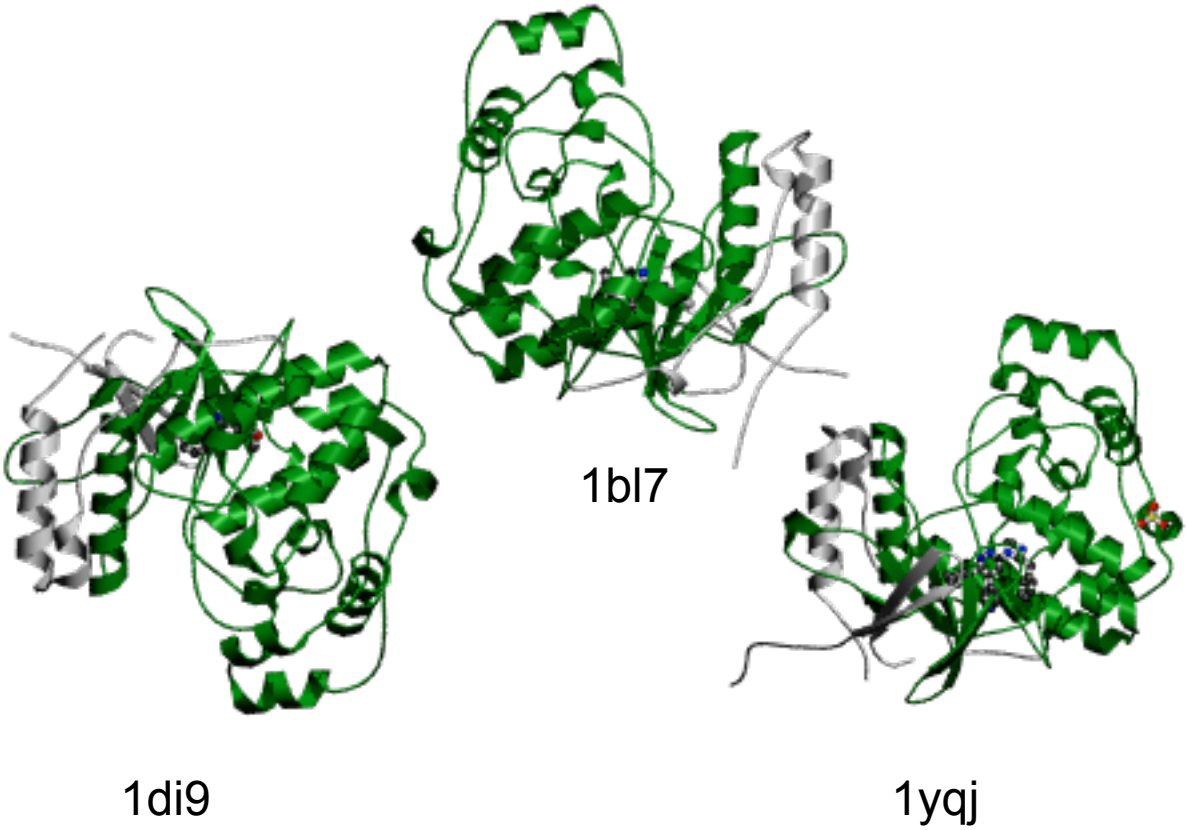
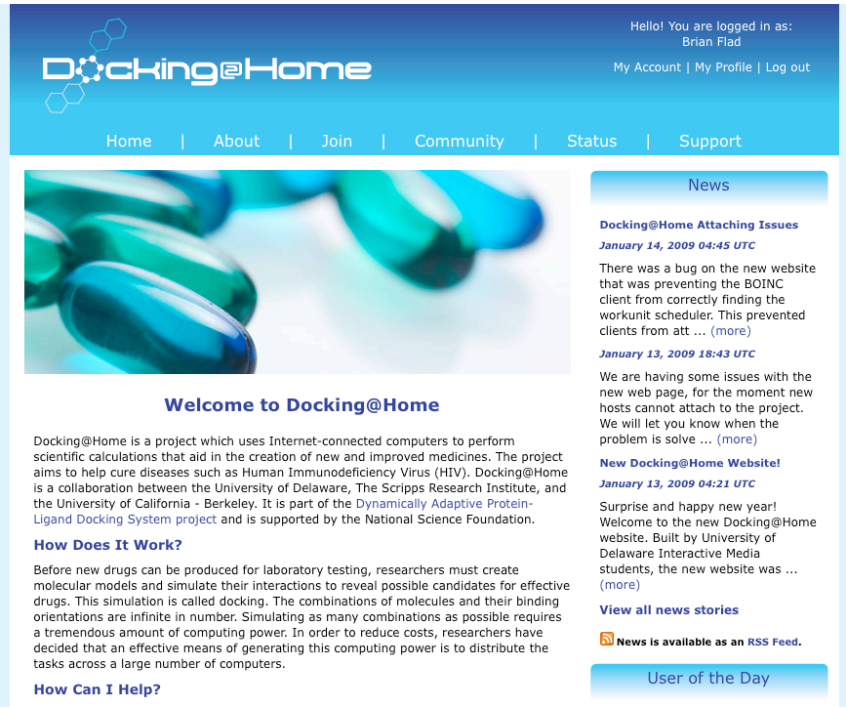


Figure 4: Three p38 alpha complexes with their ligand docked in the protein

The new Docking@Home Website

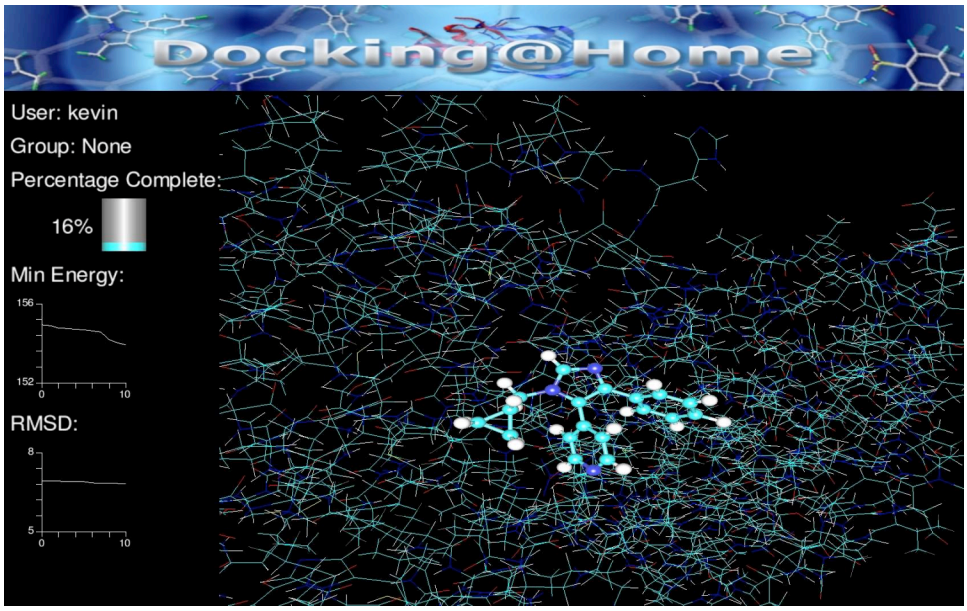
Students from the University of Delaware's Interactive Media minor (Allison Casey, Brian Flad, Brianna Higgins, Lucas Newman, and Wybens Titus) dedicated their fall semester to creating a new website structure and design for the Docking@Home website. After analyzing the results from volunteers and our team surveys, they set out to create a highly organized, simplified, and unified layout. By the end of the semester, a fully operable proof-of-concept website was completed. Since then, Brian has been working to implement the teams' work into the original content and BOINC supplied code using PHP and MySQL. He is also working on rewriting some parts of the website to be more user friendly. Many thanks to all of the volunteers who have been providing feedback throughout the project.



The new Docking@Home screensaver

The Docking@Home screensaver and graphical application serve as a visual confirmation of the volunteer's contribution to the science behind Docking@Home. Volunteers can view and manipulate 3D representations of the protein and ligand structures. These structures can be viewed separately as well as in their current docking configuration with respect to the computation being performed. Several visualization styles are also provided to aid in visualizing the physical structure of the molecules. Several metrics related to the computation are available to the volunteers so that they may monitor the job's progress.

The Docking@Home graphical application and screensaver was developed by Robert Keller and Kevin Kreiser at the University of Delaware. The application is written in OpenGL and is compatible with both Windows and Linux operating systems. Through careful planning and discussion, the development team at Docking@Home decided that an informative interactive graphical application would be well suited to achieving our goals of educating our volunteers about the docking procedure. We hope that as a result our volunteers will feel engaged in the science they are helping to progress.



Events and publications

Our events

David P. Anderson at the 1st East Coast BOINC meeting

Our team organized the first-ever East Coast meeting of BOINC on Friday, Aug. 29, at the University of Delaware. David P. Anderson, the pioneer of the volunteer computing paradigm and Berkeley Open Infrastructure for Network Computing (BOINC) of the University of California at Berkeley, spoke at the event. Trilce, Mark Somers (University of Leiden), Michael P. Cummings (University of Maryland - College Park), Jack Shultz (National Academies) and Andrew Gillette (The Foundation for Computational Learning & Science), and Eric Myers (LIGO Hanford Observatory) were invited speakers. We had 45 attendees; some of them from the University of Delaware and some from other universities and other states. You can find more about the event at:

http://gcl.cis.udel.edu/EastCoast08/

"Our team organized the first-ever East Coast BOINC meeting on August 2008"

Our recent publications

M. Taufer, R.S. Armen, J. Chen, P.J. Teller, and C.L. Brooks III: Computational Multi-Scale Modeling in Protein-Ligand Docking. IEEE Engineering in Medicine and Biology Magazine, 2008 (In Press).
More about this paper: This paper summarizes the main scientific ideas beyond D@H.

T. Estrada, O. Fuentes, and M. Taufer: A Distributed Evolutionary Method to Design Scheduling Policies for Volunteer Computing. ACM SIGMETRICS Performance Evaluation Review Journal, Volume: 36, Issue: 3, Pages: 40 – 49, December 2008.
More about this paper: Trilce is always looking for better ways to distribute the computation among the volunteers. She presented part of this paper in a conference last May. This paper looks at scheduling policies for the distribution of jobs supported by genetic algorithms.

Our recent posters

We presented D@H in posters at three important events last year. First in May at the MASPLAS workshop in Princeton, then in November at the research booth of the University of Delaware at SC'08, and finally in December at the Computer Science Day at our department in which we celebrate the successes of our students in research. Trilce and Kevin had a lot of questions from very excited students discovering D@H.

Docking@Home history in a nutshell

"The ideas behind D@H were born in 2004 at the Scripps Research Institute. Since then, the project has grown.
This is the story of D@H in a nutshell."

Spring 2003: Michela Taufer works on a docking algorithm using CHARMM with Mike Crowley and Dan Price at the Scripps Research Institute under the supervision of Charlie Brooks. The algorithm is finished in fall and the team publish a paper titled: "Study of an Accurate and Fast Protein-Ligand Docking Algorithm based on Molecular Dynamics". The main idea was to use idle cycles for protein-ligand docking.

2004: Michela put aside the docking idea to implement Predictor@Home (the first BOINC project). With Chahm An and Charlie Brooks, she took part in CASP 2004. The team was invited to present the idea of using volunteer's computers at the CASP meeting in Gaeta, Italy in December 2004.

Chronology of D@H

2005: Michela moves to the University of Texas at El Paso (UTEP) as an assistant professor and starts designing Docking@Home based on her initial idea in the paper published in 2004.

September 2005: The National Science Foundation awards funding for developing D@H to Michela and a team of other three investigators. With this valuable financial support, the D@H adventure can finally start!

September 2006: D@H is in alpha phase. A lot of challenges for Andre, Richard, Trilce, Karina, and Michela, the D@H team at UTEP.

October 2006: Cori and Atomic Booty design the D@H logo for our old Website. Great logo!

November 2006: D@H at the Supercomputing 2006 conference. Posters of the project are presented at the UTEP booth. A lot of questions and interest!

June 2007: The D@H team moves to the University of Delaware. The project is suspended for relocation. The time goes by too slowly without D@H!

November 2007: D@H at the Supercomputing 2007 conference. The project is suspended but we still present our work at the booth of the University of Delaware. A lot of people encourage us to get back on our legs!

December 2007: New servers and new start for D@H in alpha phase. Jason helps a lot with the installation.

September 2008: Docking@Home opens the creation of new accounts to all and moves to beta phase. Finally we are crunching science! Our first target is the HIV protease. Kudos to Trilce for her great effort during the summer!

November 2008: D@H starts the distribution of a new protein, the Trypsin. Roger is monitoring our results from U. Michigan and, at UDel, Ramie joins us to review the science.

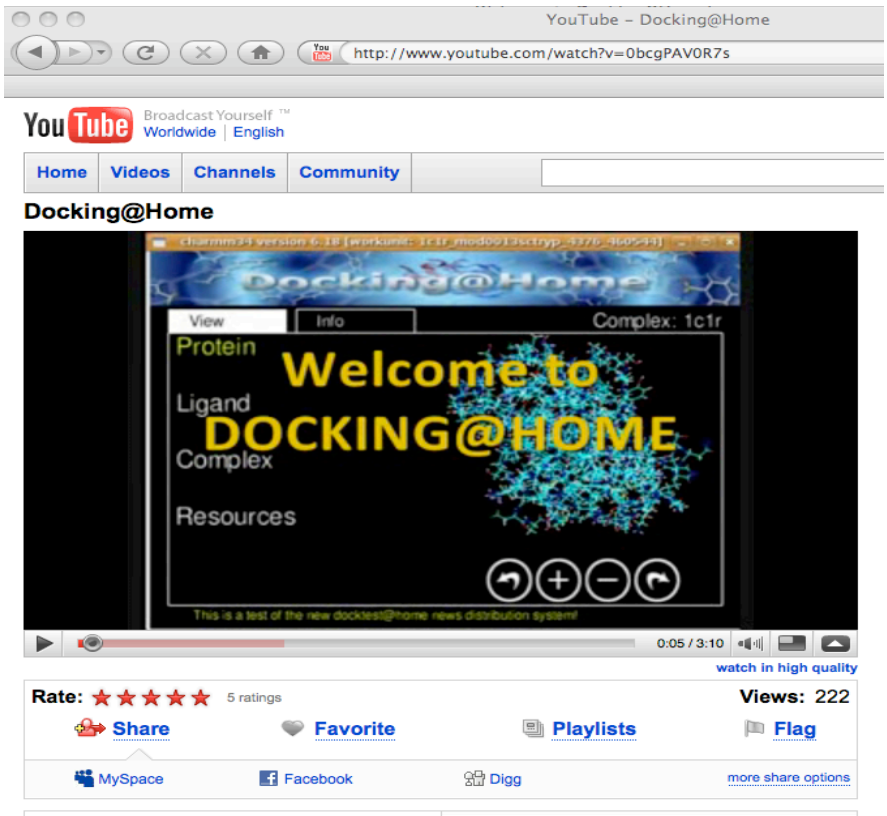
Chronology of D@H

November 2008: D@H finally has a screensaver. Look at our YouTube video: **http://www.youtube.com/watch?v=0bcgPAV0R7s.** Kudos to Kevin and Robert for the great screensaver and the professional YouTube voice-over.

November 2008: D@H at the Supercomputing 2008 conference. Yes, we are still crunching, and this time we crunch science!

January 2009: New Year and new Webpage. Many many kudos to Brian and his team for this excellent piece of art!

January 2009: A new protein is ready for D@H, the p38 alpha!!!



Our future goals

Yes, a lot of things have happened in the past year and we still have a lot in store for D@H and our volunteers. First of all, we have a new docking algorithm with more accurate representation of the solvent in which the ligands dock in the protein. We are testing the algorithm in standalone and we plan to have it ready for our volunteers' machines in March.

We are also working on methods to cluster all the results we are getting from you and summarize the science embedded in hundreds of thousands of files in an easy and clear way for Roger and Ramie (our scientists).

Our volunteers are always in our thoughts and we will give you some more fun with a new interactive screensaver. A first prototype will be available at the end of Spring 2009.

Stay tuned with D@H!

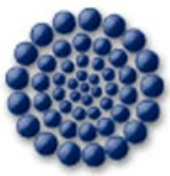
"D@H is looking forward to a brighter future with new proteins, new docking algorithms, and new scientific contributions.
Our volunteers are our *colonna portante* (in Italian it means cornerstone)"

Supporting Docking@Home

Some of you asked us how you can help D@H. The easier way is to crunch our science. D@H has been growing so fast because of our committed (and sometimes patient) volunteers. At the same time, D@H owes a lot to talented students at the University of Delaware who spend nights and weekends monitoring the system and working on new features. So far we have been very lucky because the National Science Foundation (NSF) supported us. Our NSF funding will end on August 31, 2009. Since we are a not for-profit organization and no private companies provide us with funds to pay the students working on the project, we must continuously search for new funds. The support for a graduate student is $25,000 per year for a stipend and university benefits. We are currently applying for new funds from the National Science Foundation and other national agencies: we hope we will be able to receive something in this highly competitive environment. If you want to know more how to support a student working on D@H financially, get in touch with Michela Taufer at **taufer@udel.edu**.

Our Sponsors

NSF OCI #0802650



CONACYT

This Newsletter was written by …

Michela Taufer
Trilce Estrada
Brian Flad
Kevin Kreiser
and with the feedback and suggestions of the other members of the Global Computing Lab at the University of Delaware.

http://gcl.cis.udel.edu

For questions contact us at: **dockingadmin@cis.udel.edu**