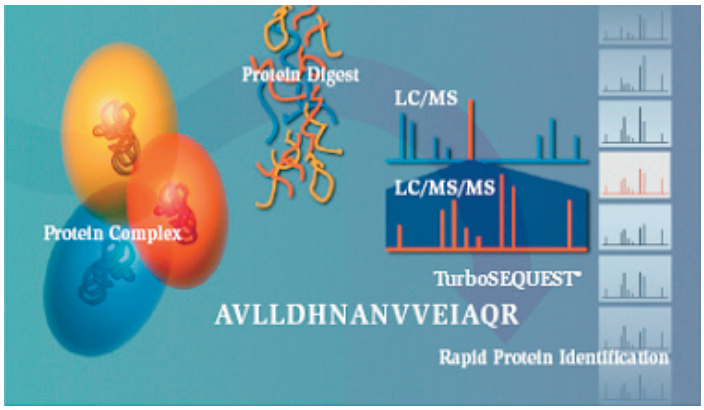
Center for High-Performance Computing @ the University of Utah

News

University of Utah Spring, 2005



Article

Clinical Proteomics at ARUP Laboratories

by Kojo S.J. Elenitoba-Johnson M.D.
Associate Professor, Department of Pathology, University of Utah School of Medicine
with David K. Crockett M.S.
Research Scientist, Clinical Proteomics, ARUP Institute for Clinical and Experimental Pathology

Completion of the human genome project has focused recent scientific attention on the definition of gene products (proteins) that function as the effectors of the genetic code. The term "proteome" has been introduced to describe the protein complement of the genome, and "proteomics" is defined as the comprehensive and large scale analysis of the protein properties of cells, tissues or organisms. In humans, proteomics is used to study changes in protein expression, to elucidate protein-protein interactions, or to obtain an integrated "global" view of normal cellular and disease processes.

The single most important technological development in the large-scale analysis of proteins is mass spectrometry. Ion trap tandem mass spectrometry (MS/MS) isolates a target ion, known as a "precursor" ion, by ejecting all interfering ions from the ion trap. The ions of interest remaining in the trap are then fragmented, producing an MS/MS spectrum. This spectrum becomes a structural "fingerprint" that is reproducible and uniquely characteristic for that given peptide ion. The data produced in such an MS/MS experiment yields peptide fragmentation patterns which software can analyze for information about a protein's composition, structure, and posttranslational modifications.

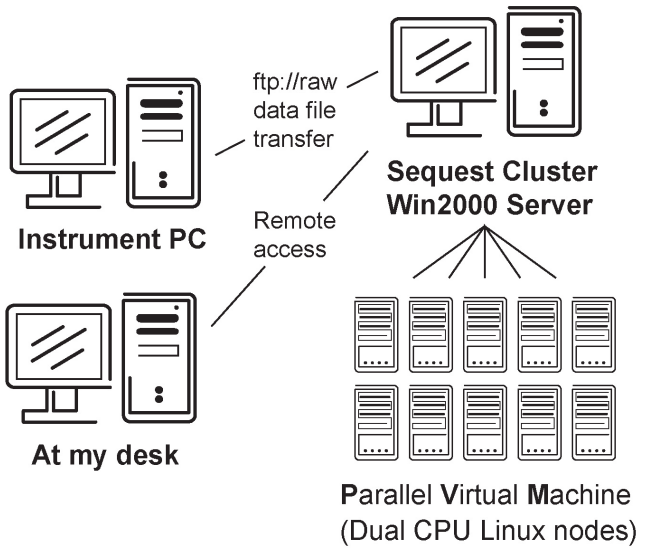
The ARUP Clinical Proteomics Group leverages expertise in pathology, protein biology, informatics and mass spectrometry. We currently study secreted proteins in a number of lymphoma models, which will eventually lead to the development of serum-based testing for diagnosis of cancer. Employing state-of-the-art methodologies in proteomics and mass spectrometry, we generate thousands of peptide sequencing events per week. Analysis of a single sample may contain as many as 4,000 MS acquired spectra. A typical experiment can yield tens of thousands of MS spectra (sequencing attempts) ready for protein database searching.

Previous limitations of computational power and manual verification of "close calls" for protein identification resulted in data analysis for each experiment taking 2 - 3 months. Similar to earlier DNA work, one obvious solution to this problem was to increase the "CPU horsepower" used when database searching. The following table gives a simple example of the power of parallel computing for protein identification.

CPU'S	SEARCH TIMES
1	19.4 hours
8	3.7 hours
16	1.1 hour

Data file: npm-alk_scx08.raw (23,407 spectra); Database: NCBI nr.fasta (2.1 million proteins)

SEQUEST Cluster® (Thermo, San Jose, CA) offers a scalable solution for computing the large volumes of data typical in most proteomics analyses using LC/MS. Protein database search times are dramatically reduced by harnessing the power of several processing units in one, increasing throughput and saving valuable analysis time. By using parallel computing our data analysis time has improved to a few days per experiment, as opposed to a few months. Quality of final protein identifications has also improved with algorithms for removing false positive search results.



Our strategic alliance with the Center for High Performance Computing (CHPC) gives ARUP's Clinical Proteomics Group access to more than $2.5 million dollars of existing computer hardware, plus expertise in parallel computing. This not only saved ARUP the cost and labor of building our own computer cluster, more importantly with CHPC's expertise, troubleshooting during the implementation of SEQUEST Cluster® was minimal.

Early experiments identified more than one thousand unique proteins using the parallel computing resources at CHPC. Numerous proteins in functional categories such as cell adhesion, migration, signaling molecules, and stress response that were not previously known in lymphoma were identified and may serve as novel disease markers and provide insight into the pathogenesis of lymphoma. This demonstrates the utility of currently available bioinformatics tools for the robust identification and annotation for large numbers of proteins in a batch-wise fashion.

In recent experiments, a total of 368 proteins were identified and fully annotated in a cancer cell line, with 124 of those proteins showing changes in quantitated expression levels. *In silico* analysis of functional groups of the over-expressed proteins included protein kinases, cytoskeletal proteins and proteins associated with cell proliferation. This study demonstrates that global proteomic consequences of disease can be studied using tandem mass spectrometry and high performance computing.

Overall, the protein identifications generated utilizing SEQUEST Cluster® have been promising. Our data has been featured in a growing number of published manuscripts specializing in cancer pathology and proteomics. The continued use of CHPC's expertise and resources in parallel computing are key to the success of ARUP's Clinical Proteomics research group.

Upcoming Presentations

CHPC has developed a series of courses to help users make the most use of CHPC resources. Our spring series begins March 24th. Please mark your calendars. These presentations are all held in the INSCC Audtitorium and begin at 1:30pm on the scheduled date:

March 24th: Overview of CHPC

March 31st: Introduction to Parallel Computing

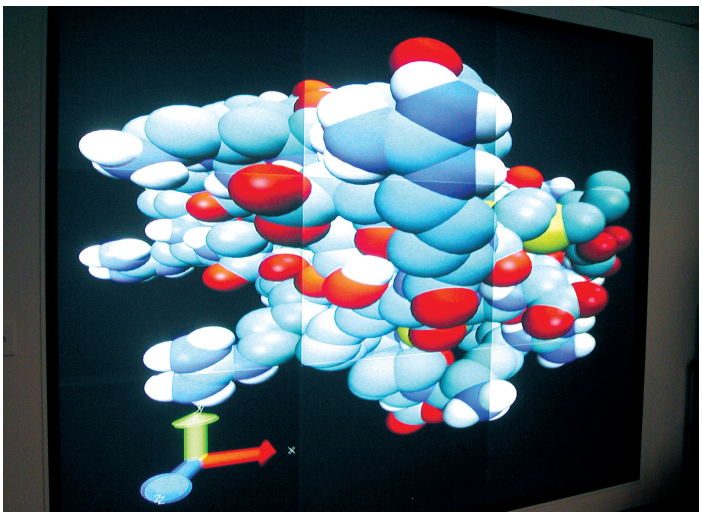
April 7th: Chemistry Packages at CHPC

April 21st: Using Gaussian03 and Gaussview

April 28th: Introduction to Programming with MPI

May 5th: Debugging with Totalview

Slides from CHPC's presentations are archived on the CHPC web site. You may access them at any time by going to http://www.chpc.utah.edu/docs/presentations/ and selecting the name of the presentation either from the menu tree or the presentation list in the central content area.



Article

Skyline Arch: CHPC's New Visualization Cluster

by Sam Liston

Digital Communication & Visualization, Center for High Performance Computing, University of Utah

CHPC's new visualization cluster is complete and ready for use. Skyline Arch, which consists of ten dual Opteron nodes driving 18 Sanyo LCD projectors, is capable of displaying stereo images as large as 3072 x 2304.

Skyline Arch is a distributed visualization cluster. It is powered by Chromium, a continuation of the Stanford WireGL project, which runs "beneath" an application and creates a "tiled" display from its graphics information. The head node (application node) runs an OpenGL application. The OpenGL calls — normally directed to the local graphics card — are intercepted by Chromium and redirected via some "interconnect" (i.e. TCP/IP, Myrinet, Infiniband) to the backend nodes (client nodes) where the calls are interpreted, rendered by the remote graphics card and displayed.

Each client node is aware of its area of responsibility. It knows the dimension of its render area as well as its location in the overall display. As objects get passed from one portion of the display to another, each client node must know when an object is leaving or coming to its portion of the display. This intercommunication is crucial in order to keep the movement of an object smooth as it transitions across a display border. This constant intercommunication can also hinder performance. As the number and complexity of the displayed objects increases, latency can quickly become an issue. Because of this, Myrinet is used for intercommunication. The low-latency, high-bandwidth solution greatly improves the display wall's ability to refresh displays fast enough and

in synchronization so they are not discernible by the eye.

The stereographic aspect of Skyline is achieved using a century old technology made famous by the "stereoscope" of the 19th century and cheesy 3D horror movies of the 20th century. The effect is created by rendering two separate views of the same object, offset slightly in their orientation from one another: one view for the left eye and one for the right. There are two projectors for each portion of the display. Each projector is outfitted with a polarized lens. These lenses are offset 90 degrees from one another. When the stereo display is viewed through a pair of polarized glasses, the left eye sees only the image rendered for the left and the right eye only sees the image rendered for the right. Though archaic, this method of viewing stereo images is quite effective; it is easy on the eyes and is reasonably priced.

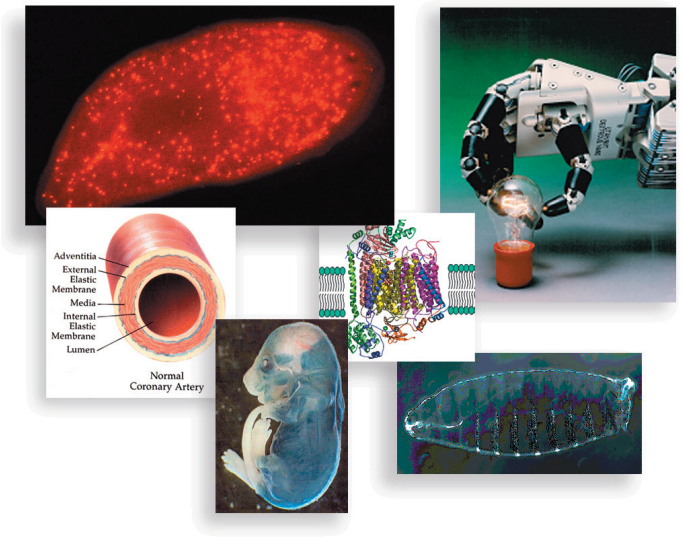
Several applications have been tested on Skyline Arch. For instance, Visual Molecular Dynamics (VMD) has been thoroughly tested. With VMD we have been able to display, in stereo, structures on the order of 350,000 atoms. NCSA Pixel Blaster has also been tested. This application allows the viewing of high-resolution images and sequences of images. Most of the standard image formats are supported (JPEG, GIF, TIFF, etc). Paraview, a graphical front-end to the Visualization Toolkit (VTK), has also been tested. Paraview is primarily used for volume rendering.

Almost any OpenGL based application will run on Skyline Arch. This is a major benefit of using Chromium to power the display: it is quite flexible. The exception is that not all visualization applications understand "stereo." In its current state, Skyline Arch may only be able to display a visualization in two dimensions. We are working on ways to remedy this, specifically to force applications through Chromium to output stereo images. This is still in the testing stage, both in the development of Chromium and our display wall.

If you have a visualization application you would like to use on Skyline Arch, please contact Sam Liston (stliston@chpc.utah.edu).

FYI

The formal inauguration of Michael K. Young as the 14th president of the University of Utah will take place Friday, April 15, 2005. Celebratory events will be held that day and for several days prior. Please keep Friday, April 15th open so that all members of the University family can participate in this special ceremony. Additional information will be made available as the inauguration approaches.



Report

On the Scene: Research Posters on the Hill, 2005

by Robert McDermott

Staff Scientist, Visualization Group, Center for High Performance Computing, University of Utah

This year's Posters on the Hill event took place on January 20th. Due to the Capitol Rotunda renovation project, the event was relocated to the auditorium of the State Office Building. Jill Bader of the Undergraduate Research Opportunities Program solicited UROP students for the majority portion of participants. The Offices of the Dean of Science and Dean of Engineering also contributed students to participate in the event.

The Offices of Governmental Affairs and of the Vice President of Research contributed to the success of the event. This year, Julio Facelli, Director of the Center for High Performance Computing, generously contributed an intern, Iris Boanta, who helped with assorted tasks during the day of the event.

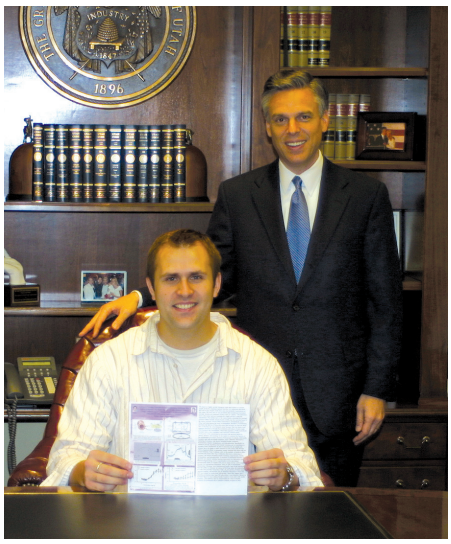
This year, 32 posters were presented to the legislators by 35 undergraduate student researchers. The change in venue for the event mandated a number of changes to the presentation of the posters. Due to a reduction in space, the posters were designed to be more vertical than horizontal and were spaced within a few inches of each other. These changes received some unexpected support at the event. Not only were the vertical poster designs easier to read, but the grouping of the posters provided a more intimate setting for the legislators to talk with the students.

A major concern for this year was that fewer legislators would be willing to walk the distance to the State Office Building auditorium to see the students and their posters. This turned out to be a non-issue: legislators stood up in their respective chambers to announce that the students with their posters were in the State Office Building auditorium. With this support there was an excellent showing of legislators, comparable to the showing in past years when the event was held in the Capitol Rotunda.

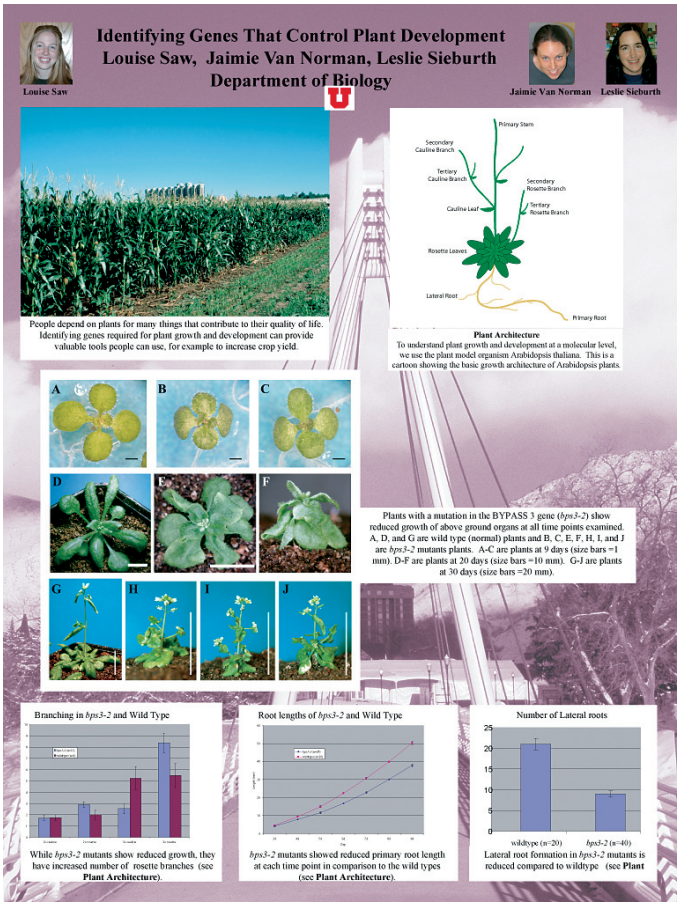
One of our students Jeff Johnson, through his persis-

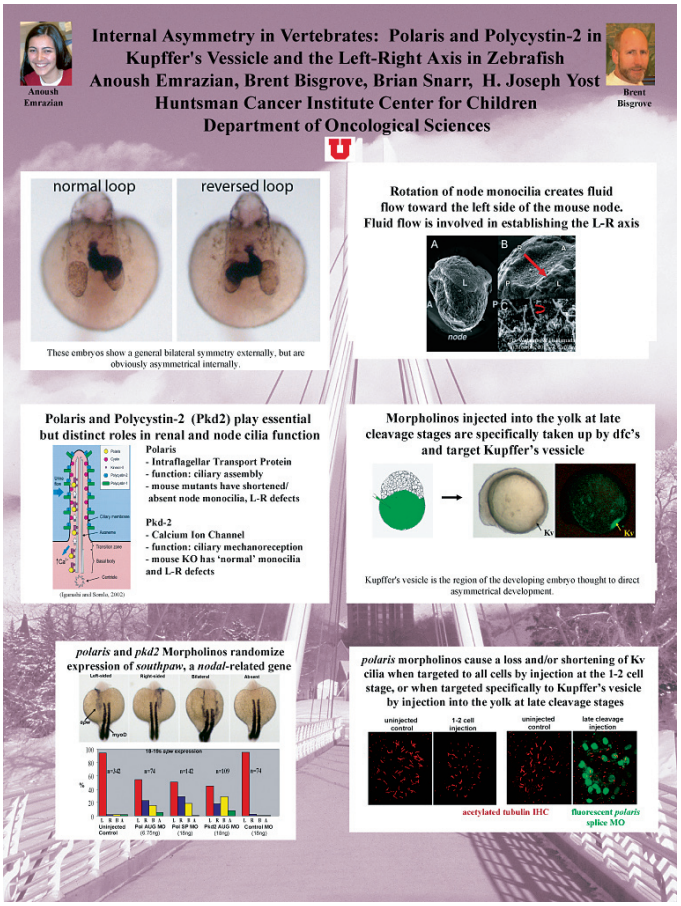
tence and the help of a friend working in the Utah State Governor's Office, managed to have his photo taken with Governor John Huntsman.

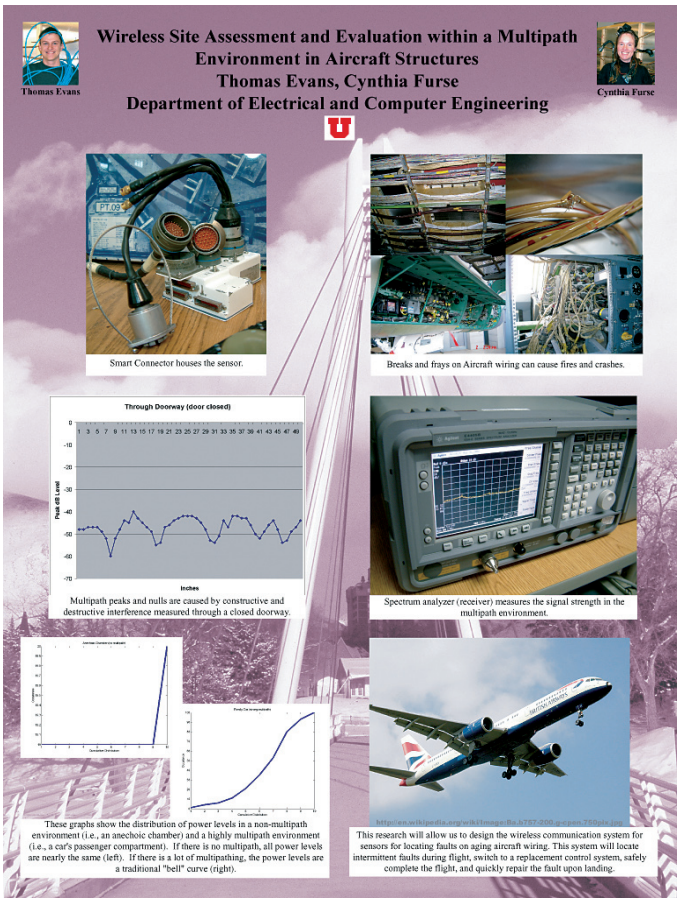
The five posters included with this article are representative of not only the high quality but also the diversity of research represented by this year's participants. Students are the fabric of this lobbying effort. Providing them with visually engaging posters to stand by when talking with their state legislators has proven once again to be a successful combination.



To see images of all the posters from this year's Posters on the Hill event, go to the CHPC web site under "Docs|Research" (http://www.chpc.utah.edu/docs/research") and select the "Posters on the Hill - 2005" link. Images from previous Posters on the Hill events are also available.







Recent Publications

Bazterra, V. E., M. Cuma, et al. (2005). "A general framework to understand parallel performance in heterogeneous clusters: analysis of a new adaptive parallel genetic algorithm." Journal of Parallel and Distributed Computing 65: 48-57.

Cheng, W. Y. Y. and W. J. Steenburgh (2004). "Evaluation of surface sensible weather forecasts by the WRF and Eta models over the Western United States." Weather and Forecasting Submitted.

Di Fiori, N., Orendt, A.M., Caputo, M.C., Ferraro, M.B., and Facelli, J.C. (2004). "Modeling solid-state effects on NMR chemical shifts using electrostatic models." Magnetic Resonance in Chemistry 42: 41-47.

Espinal, J. F., F. Mondragon, et al. (2004). "Density Functional Theory Study of Carbon-H2O Reactions during Gasification with Steam." American Chemical Society, Division of Fuel Chemistry 49(822).

Espinal, J. F., A. Montaya, et al. (2004). "A DFT Study of Interaction of Carbon Monoxide with Carbonaceous Materials." Journal of Physical Chemistry B 108: 1003-1008.

Facelli, J. C., Sefzik, T.H., et al. (2005). "Modeling NMR Chemical Shift: A Survey of Density Functional Theory Approaches for Calculating Tensor Properties." Journal of Physical Chemistry 109: 1180-1187.

Freedman, H. and T. H. Truong (2004). "An Application of Coupled Referenced Interaction Site Model (RISM)/ Molecular Dynamics (MD) Method to the Conformational Analysis of the Alanine Dipeptide." Journal of Physical Chemistry (In Press).

Freedman, H. and T. H. Truong (2004). "A Coupled reference interaction site model (RISM)/molecular dynamics (MD) study of potential mean force of the SN2Cl + CH3Cl reaction." Journal of Physical Chemistry (Submitted).

Freedman, H. and T. H. Truong (2004). "A Study of the Tautomic Equilibria of 2-hydroxypyridine/2-oxopyridine and of Cystosine in Water Using the Coupled Reference Interaction Site Model (RISM)/Molecular Dynamics (MD) Approach." Journal of Physical Chemistry (In Press).

Freedman, H. and T. N. Truong (2004). "Coupled Reference interaction Site Model/Simulation Approach for Thermochemistry of Solvation: Theory and Prospects." Journal of Chemical Physics 121(2187).

Hart, K. A., W. J. Steenburgh, et al. (2004). "Model forecast

improvements with decreased horizontal grid spacing over fine-scale Intermountain orography during the 2002 Olympic Winter Games." Weather and Forecasting (Submitted).

Huang, M. H., M. Cuma, et al. (2004). "Bending of Nanoscale Thin Si Film Induced by Growth of Ge Islands: Hut vs. Dome." MRS Proceeding 791(Q6.4): 1-6.

Jenkins, M. A. (2004). "Investigating the Haines Index using parallel model theory." International Journal of Wildland Fire (Pending).

Jungsuttiwong, S., J. Limtrakul, et al. (2004). "Theoretical study of Modes of Adsorption of Water Dimer on H-ZSM-5 and H-Faujasite Zeolites." Journal of Physical Chemistry (Submitted).

Ka, B. J. and G. A. Voth (2004). "Combining the Semiclassical Initial Value Representation with Centroid Dynamics." J. Phys. Chem B 108(21): 6883-6892.

Krongpracha, P., P. Treesukol, et al. (2004). "A Theoretical Study of Adsorption of Carbon Monoxide and Nitrogen Oxide on Ag-ZSM-5 Zeolite." Journal of Physical Chemistry (Submitted).

Limtrakul, J., C. Inntam, et al. (2004). "Density Functional Theory Study of the Ethylene Epoxidation over Ti-substituted Silicalite (TS-1): An Application of Cluster and Embedded Cluster Methods." Journal of Molecular Catalysis A: Chemical 207(139).

Lu, G.-H., M. Cuma, et al. (2004). "Quantitative understanding of strain stabilization of Ge/Si(105) surface from first-principles." Physical Review Letters (Submitted).

Lu, G.-H., M. Huang, et al. (2004). "Relative stability of Si surfaces: a first principles study." Physical Review B (Submitted).

Nguyen, H. N., S. Zhang, et al. (2004). "Direct ab initio Dynamics Studies of the Reactions of HNO with H and OH radicals." Chemical Physical Letters 388(94).

Pala, R. and F. Liu (2004). "Determining the Adsorptive and Catalytic Properties of Strained Metal Surfaces using Adsortion-induced Stress." Journal of Chemical Physics 120(7720).

Reichler, T., Kushner, P.J., and Polvani, L.M. (2004). "Response of the troposphere-stratosphere system to impulsive tropospheric forcing." Journal of the Atmospheric Sciences (Submitted).

Truong, T. N. (2004). "Computational Science and Engineering Online: An Integrated Web-based Environment for Multi-scale Modeling of Complex Reaction Systems. (Proceedings of the 2nd International Conference on Foundations of Molecular Modeling and Simulations)." Molecular Physics 102(353).

Violi, A., T. N. Truong, et al. (2004). "Kinetics of Hydrogen Abstraction Reactions From Polycyclic Aromatic Hydrocarbons by H atoms." Journal of Physical Chemistry A 108(4846).

Wang, Y. and T. N. Truong (2004). "Theoretical Study of Adsorption of Water Dimer on the Perfect MgO (100) Surface: Molecular Adsorption versus Dissociative Chemisorption." Journal of Physical Chemistry B 108: 3289-3294.

Yan, B., F. Liu, et al. (2004). "Local strain-mediated chemical potential control of quantum dot self-organization in heteroepitaxy." Physical Review Letters 92(025502).

Yan, B., P. Zhang, et al. (2004). "A novel mechanism for self-organization of semiconductor nanocrystals by selective surface faceting process." Physical Review Letters (Submitted).

Zang, J., A. Treibergs, et al. (2004). "Geometry Constant Defining Shape Transitions of Carbon Nanotubes under Pressure." Physical Review Letters 92(105501).

CHPC Security Policies

Please read and comply with the University of Utah Information Resources Policies, particularly sec. C and D.

CHPC does not allow clear text passwords when accessing our systems. We require the use of Secure Shell (SSH).

You may not share your account with anyone under any circumstances.

Do not leave your terminal unattended while you are logged in to your account.

Do not introduce classified or sensitive work on CHPC systems.

Protect your password and follow the password policies outlined at http://www.chpc.utah.edu/docs/policies.

Do not try to break passwords, tamper with system files, look into anyone else's directories, or otherwise abuse the trust implicit in your account.

Do not inspect, modify, distribute, or copy privileged data or software without proper authorization, or attempt to do so.

If you suspect a security problem, report it promptly to CHPC's Help Desk. Phone: (801) 971-3442 email: problems@chpc.utah.edu. If your concerns are an emergency during non-University working hours, please contact the campus help desk at 581-4000.

CHPC Staff Directory

Administrative Staff	Title	Phone	Email	Location
Julio Facelli	Director	556-2426	Julio.Facelli@utah.edu	410 INSCC
Julia Harrison	Associate Director	652-0019	julia@chpc.utah.edu	430 INSCC
Guy Adams	Assistant Director, Systems	554-0125	gadams@chpc.utah.edu	424 INSCC
Joe Breen	Assistant Director, Networking	550-9172	jbreen@chpc.utah.edu	426 INSCC
DeeAnn Raynor	Administrative Officer	581-5253	dee@chpc.utah.edu	412 INSCC
Victoria Volcik	Administrative Assistant	585-3791	vicky@chpc.utah.edu	405-2 INSCC
Teresa Hennigan	Administrative Assistant	581-6440	teresa@chpc.utah.edu	405-3 INSCC

Scientific Staff	Expertise	Phone	Email	Location
James Agutter	Information Visualization	581-8779	agutterja@arch.utah.edu	235 AAC
Thomas Cheatham III	Biomolecular Modeling	587-9652	cheatham@chpc.utah.edu	306 INSCC
Martin Cuma	Scientific Applications	587-7770	mcuma@chpc.utah.edu	418 INSCC
Byron L. Davis	Statistics	585-5604	byron@chpc.utah.edu	416 INSCC
Julio Facelli	Molecular Sciences	556-2426	Julio.Facelli@utah.edu	410 INSCC
Stefano Foresti	Information Visualization	581-3173	stefano@chpc.utah.edu	322 INSCC
Robert McDermott	Visualization	581-4370	mcdermott@chpc.utah.edu	420 INSCC
Anita Orendt	Molecular Sciences	231-2762	orendt@chpc.utah.edu	422 INSCC
Alun Thomas	Bioinformatics	587-9309	alun@gene.pi.med.utah.edu	Research Park

Systems/Network Staff	Title	Phone	Email	Location
Irvin Allen	System Administrator	231-3194	iallen@chpc.utah.edu	405-40 INSCC
Wayne Bradford	System Administrator	243-8655	wayne.bradford@chpc.utah.edu	405-41 INSCC
Erik Brown	System Administrator	824-4996	erik@chpc.utah.edu	405-29 INSCC
Joe Clyde	Network Operations Engineer	558-7661	joe.clyde@chpc.utah.edu	405-38 INSCC
Brian Haymore	Lead, Comp. Cluster Admin.	558-1150	brian@chpc.utah.edu	428 INSCC
Samuel T. Liston	Digital Communication & Visualization	232-6932	stliston@chpc.utah.edu	405-30 INSCC
Jimmy Miklavcic	Multimedia, Telematic & Digital Communication	585-9335	jhm@chpc.utah.edu	296 INSCC
Ron Price	Software Engineer & Grid Architect	560-2305	rprice@eng.utah.edu	405-4 INSCC
David Richardson	Computer Technician	550-3788	drr@chpc.utah.edu	405-23 INSCC
Steve Smith	System Administration	581-7552	steve@chpc.utah.edu	405-25 INSCC
Eli Stair	System Administrator	558-3099	eli@chpc.utah.edu	405-39 INSCC
Matthew Thorley	Network Assistant	560-3438	ruach@chpc.utah.edu	405-20 INSCC
Neal Todd	System Administrator	259-3495	neal@chpc.utah.edu	405-31 INSCC
Alan Wisniewski	Network Engineer	580-5835	quantix@chpc.utah.edu	405-21 INSCC

User Services Staff	Title	Phone	Email	Location
Iris Boanta	Technical Assistant	N/A	iris@chpc.utah.edu	405-10 INSCC
Jason Duhaine	Systems Assistant	N/A	jason@chpc.utah.edu	405-28 INSCC
Shawn Lyons	Network Assistant	N/A	slyons@chpc.utah.edu	405-22 INSCC
Beth Miklavcic	Multimedia Design, Digital Video	585-1067	bam@chpc.utah.edu	115 INSCC
Liza Newren	Technical Assistant	N/A	liza@chpc.utah.edu	405-9 INSCC
Erik Ratcliffe	Graphic & Web Design	N/A	erat@chpc.utah.edu	405-13 INSCC

The University of Utah seeks to provide equal access to its programs, services, and activities to people with disabilities. Reasonable prior notice is needed to arrange accommodations.

UNIVERSITY OF UTAH
Center for High Performance Computing
155 South 1452 East, RM #405
SALT LAKE CITY, UT 84112-0190

Welcome to CHPC News!

If you would like to be added to our mailing list, please fill out this form and return it to:

Vicky Volcik
UNIVERSITY OF UTAH
Center For High Performance Computing
155 S 1452 E ROOM 405
SALT LAKE CITY, UT 84112-0190
FAX: (801)585-5366

(room 405 of the INSCC Building)

Name:
Phone:

Department or Affiliation:
Email:

Address:
(UofU campus or U.S. Mail)

Thank you for using our Systems!

Please help us to continue to provide you with access to cutting edge equipment.

ACKNOWLEDGEMENTS

If you use CHPC computer time or staff resources, we request that you acknowledge this in technical reports, publications, and dissertations. Here is an example of what we ask you to include in your acknowledgements:

"*A grant of computer time from the Center for High Performance Computing is gratefully acknowledged.*"

If you use Arches, please add:

"partially supported by NIH-NCRR grant # 1S10RR17214."

Please submit copies of dissertations, reports, preprints, and reprints in which the CHPC is acknowledged to: Center for High Performance Computing, 155 South 1452 East, Rm #405, University of Utah, Salt Lake City, Utah 84112-0190