UCLA-DOE Institute for Genomics and Proteomics
Final Technical Report (9/30/87-6/30/02)

I. Scope and Mission of the Institute

The UCLA-DOE Institute of Genomics and Proteomics is an organized research unit of the University of California, sponsored by the Department of Energy through the mechanism of a Cooperative Agreement. Today the Institute consists of 10 Principal Investigators and 7 Associate Members, developing and applying technologies to promote the biological and environmental missions of the Department of Energy, and 5 Core Technology Centers to sustain this work. The focus is on understanding genomes, pathways and molecular machines in organisms of interest to DOE, with special emphasis on developing enabling technologies.

Since it was founded in 1947, the UCLA-DOE Institute has adapted its mission to the research needs of DOE and its progenitor agencies as these research needs have changed. The Institute started as the AEC Laboratory of Nuclear Medicine, directed by Stafford Warren, who later became the founding Dean of the UCLA School of Medicine. In this sense, the entire UCLA medical center grew out of the precursor of our Institute. In 1963, the mission of the Institute was expanded into environmental studies by Director Ray Lunt. I became the third director in 1993, and in close consultation with David Galas and John Wooley of DOE, shifted the mission of the Institute towards genomics and proteomics. Since 1993, the Principal Investigators and Core Technology Centers are entirely new, and the Institute has separated from its former division concerned with PET imaging.

The UCLA-DOE Institute shares the space of Boyer Hall with the Molecular Biology Institute, and assumes responsibility for the operation of the main core facilities. Fig. 1 gives the organizational chart of the Institute.

Some of the benefits to the public of research carried out at the UCLA-DOE Institute include the following:

The development of publically accessible, web-based databases, including the Database of Protein Interactions, and the ProLinks database of genomicly inferred protein function linkages.

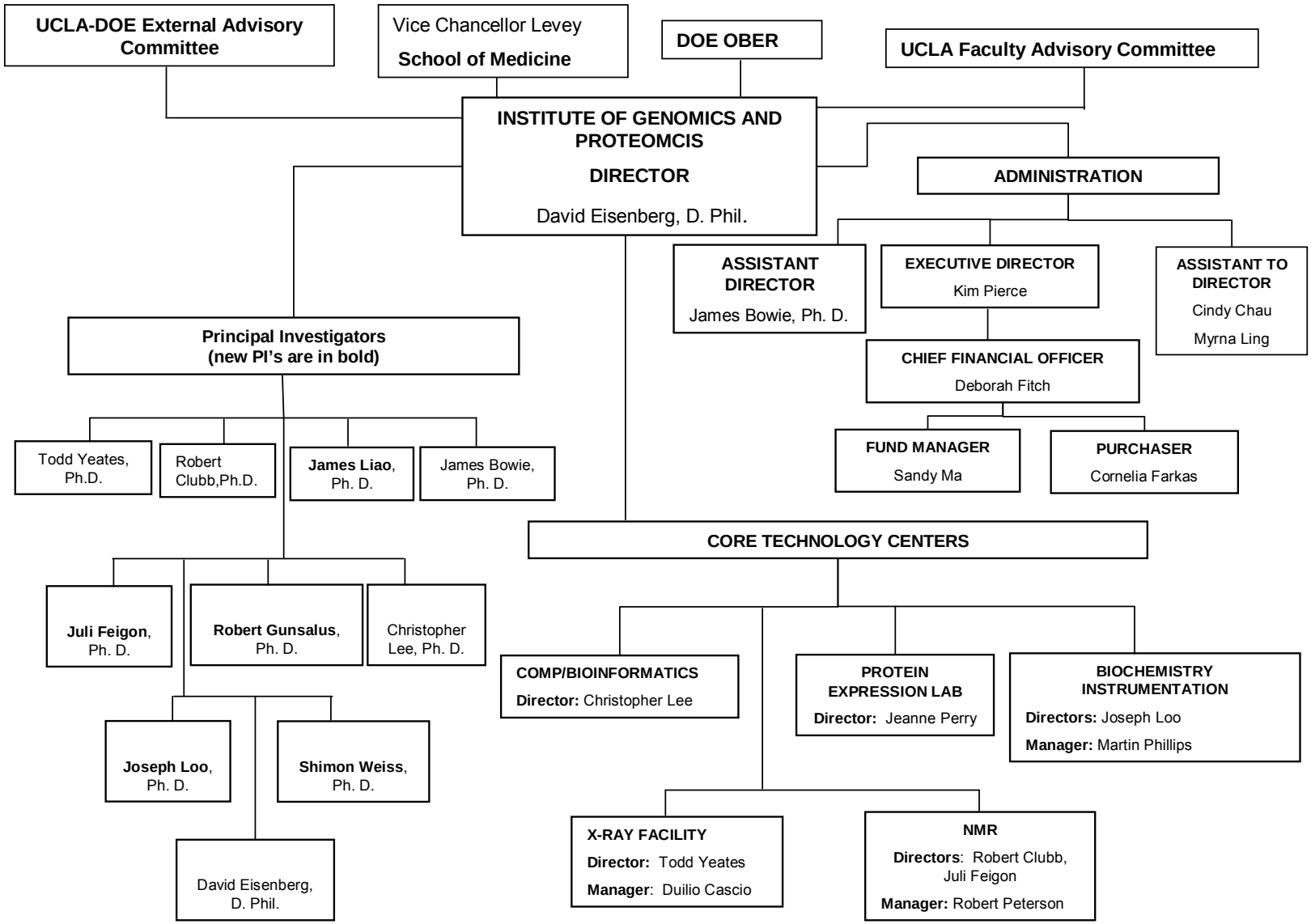
The development of publically accessible, web-based servers, including the HOTPATCH server, the ProKnow Server and the SAVEs server. All of thse are accessible from the home page of the Institute.

Advancing the science of bioenergy, in the laboratories of the Principal Investigators of the Institute, including the laboratories of Shimon Weiss, James Liao, James Bowie, Todd Yeates, Rob Gunsalus.

II. The Review of 2001 and Establishment of the External Advisory Committee

The Institute was last reviewed in 2001, when the review group and DOE staff found excellent science but inadequate focus on new DOE programs. Since then a major effort has been made to apply our new technologies to problems and organisms of greatest current interest to DOE. We have been helped in this refocus by the appointment of an External Advisory Committee which met in September 2003, and made recommendations to us and DOE staff. Figures 2 shows the personnel of the External Advisory Committee and DOE staff concerned with the Institute.

Fig. 1 UCLA-DOE INSTITUTE OF GENOMICS AND PROTEOMICS



Recommendations that came out of the 2003 Advisory Committee meeting included:

- Continued refocus on both discovering new technologies to aid the DOE program and focus on organisms of special interest to DOE
- A funding cycle of 5 years, rather than the 3 years of the current period
- The opportunity of the Institute to apply for and receive some large ticket equipment items to advance the mission, as had been part of the 1993 agreement to reorient the Institute towards genomics and proteomics.

Fig. 2 External Advisory Committee

Ruedi Aebersold
Institute for Systems Biology
1441 North 34th Street
Seattle, Washington 98103-8904
Phone: (206) 732-1204
Email: raebersold@systemsbiology.org

Michael Levitt
Professor of Structural Biology
Fairchild D109
Stanford University
Stanford, California 94305-5126
Phone: (650) 723-6800
Email: Michael.levitt@stanford.edu

Priscilla Cooper
Life Sciences Division
Lawrence Berkeley National Laboratory
One Cyclotron Road, MS 74R0157
Berkeley, California 94720
Phone (510) 486-7346
Email: pkcooper@lbl.gov

Elbert Branscomb
Joint Genome Institute
Department of Energy
2800 Mitchell Drive
Walnut Creek, California 94598
Phone: (925) 296-5608
Email: ewbranscomb@lbl.gov

Robert Stroud
Professor of Biochemistry & Biophysics
Box 2240
Genentech Hall S412C
University of California, San Francisco
San Francisco, California 94143
Phone: (415) 476-4224
Email: stroud@msg.ucsf.edu

Jim Fredrickson
Pacific Northwest National Laboratory
P.O. Box 999/MS P7-50
Richland, Washington 99352
Phone: (509) 376-7063
Email: Jim.Frederickson@pnl.gov

Michelle Buchanan
Director
Chemical Sciences Division
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN 37831
Phone: (865) 574-4986
Email: buchananmv@ornl.gov

Barbara Wold
Bren Professor of Molecular Biology
California Institute of Technology
Mail Code 156-29
Pasadena, California 91125
Phone: (626) 395-4916
Email: woldb@caltech.edu

Dr. Marvin E. Frazier, Director
Office of Biological and Environmental Research
Life Sciences Division
SC-72/Germantown Building
U. S. Department of Energy
1000 Independence Avenue, SW
Washington, DC 20584-1290
Phone: (301) 903-5468; Fax: (301) 903-8521
Email: marvin.frazier@oer.doe.gov

Dr. Roland Hirsch
Environmental Remediation Sciences Division
Office of Biological And Environmental Research
SC-75/Germantown Building
U.S. Department of Energy
1000 Independence Avenue, SW
Washington, DC 20585-1290
Phone: (301) 903-9009; Fax: (301) 903-4154
Email: roland.hirsch@science.doe.gov

David Thomassen
U.S. Department of Energy
Office of Science
Biological & Environmental Research
19901 Germantown Road
Germantown, MD 20874-1290
Phone: (301) 903-9817
Email: david.thomassen@science.dow.gov

Dr. Michael Viola, Director
Office of Biological and Environmental Research
Medical Sciences Division
SC-73 Germantown Building
1000 Independence Avenue, SW
Washington, DC 20585-1290
Phone (301) 903-3213; Fax (301) 903-0567
Email: michael.viola@science.doe.gov

III. Refocus of the UCLA-DOE Institute on Research Relevant to DOE

The refocus of the Institute has several components:

The first is the addition of five new Principal Investigators to broaden expertise, enabling better focus on DOE's mission. The new PIs and their expertise are:

Dr. Rob Gunsalus – Microbiology of DOE-mission organisms
Dr. Joseph Loo – Mass spectrometry and proteomics
Dr. James Liao – Metabolic networks and DNA microarray methods
Dr. Juli Feigon – RNAs and Protein-RNA Complexes
Dr. Shimon Weiss – FRET, Quantum dots and other single molecule methods

The second component of the refocus involves applications to organisms of special interest to DOE. Three organisms were selected with advice from DOE (Fig. 3). They are:

Rhodopseudomonas palustris
Shewanella oneidensis MP1
Methanosarcina acetivorans C2A

The third component of the refocus is to concentrate on the discovery, structures, and actions of molecular machines.

Fig. 3a

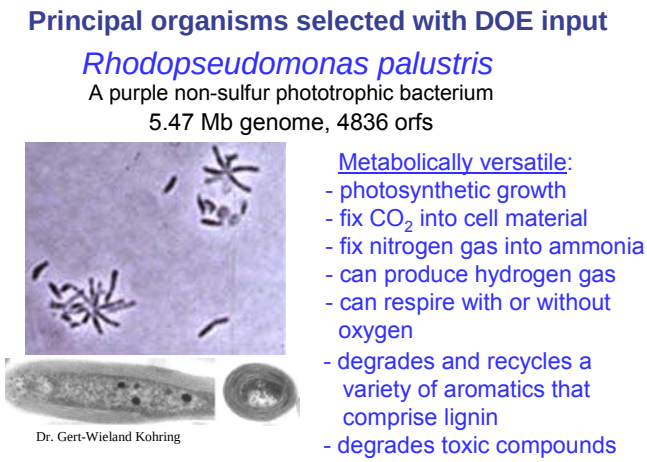


Fig. 3b

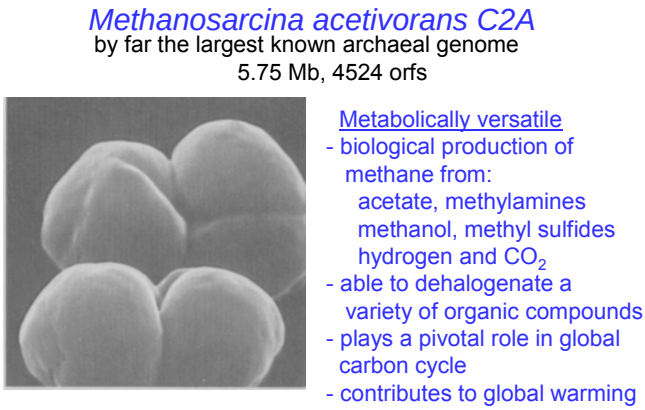
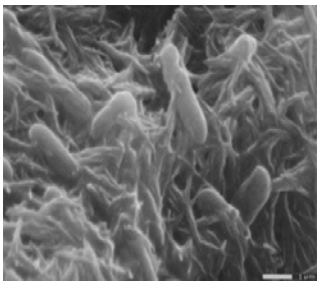


Fig. 3c

Shewanella oneidensis MP1

Widespread in soil and water environments

4.97 Mb genome, 4758 orfs



A model organism for bioremediation studies

Complex electron transport systems and metal ion-reducing capabilities.

Versatile respiration:
- oxygen
- iron
- uranium
- chromium
- non oxygen conditions

Table I summarizes current and proposed research in the Institute on these organisms

Table I Refocus of UCLA-DOE Research on DOE Missions

PI	Organism	Area	New Technologies	UCLA-DOE Collaborators
Bowie	RP, MS, S	Integral membrane proteases, membrane protein structure	Membrane protein expression and stabilization	Perry
Clubb	RP	Aromatic degradation and transcriptional regulation	Global binding site identification	Perry, Liao
Eisenberg	RP	Define protein networks / assign 3D folds	Bioinformatics / microarrays / structures of complexes	Yeates, Liao, Perry, Gunsalus, Lee, Weiss
Feigon	RP, S	Riboswitches	Riboswitch design	Gunsalus
Gunsalus	MS	Microbial communities	Microarrays	Loo, Perry, Yeates, Eisenberg
Lee	MS	Protein networks, genome evolution	Complete proteome alignments	Gunsalus, Eisenberg
Liao	RP	Regulatory networks	Microarrays and network analysis	Clubb, Eisenberg
Loo	MS	Protein interactions, modifications, microbial communities	Mass spectrometry	Gunsalus, Perry, Weiss
Weiss	S	Transcription regulation	Single molecule methods: PAID and ALEX	Perry, Eisenberg
Yeates	MS, S, RP	Genome-wide complex identification	Interaction networks by disulfide identification and logic analysis	Perry, Eisenberg, Gunsalus

S = *Shewanella*, MS = *Methanosarcina*, RP = *R. palustris*

IV. Core Technology Centers

The Institute sponsors five Core Technology Centers listed in Table II, and in the future will develop a sixth, on optical studies. Each Core Center is managed by a professional in the field. With the exception of the Protein Expression Core Technology Center, all Centers are supported from multiple sources. All of the Centers welcome other Campus users, and there are no user fees. However, the major work of each Center is on DOE research.

The Core Technology Centers depend enormously on state-of-the-art equipment. The once generous equipment sponsorship of DOE has run dry in recent years, and so we are encouraged by the recommendation that we can again apply for several large ticket equipment items. The request items and their justifications are listed in this proposal along with the descriptions of the Centers. Enhancements to the technological infrastructure of the Centers is essential to the progress of the mission. During the previous grant period we were able to enhance the NMR Center. Now equipment for the Computer/Bioinformatic Center, the X-ray Diffraction Center, and the Biochemical Instrumentation is sought. All three of these Centers are heavily used by Institute researchers. We hope to upgrade smaller items of equipment for the Protein Expression Center from other sources during the next grant period.

Table II Core Technology Centers

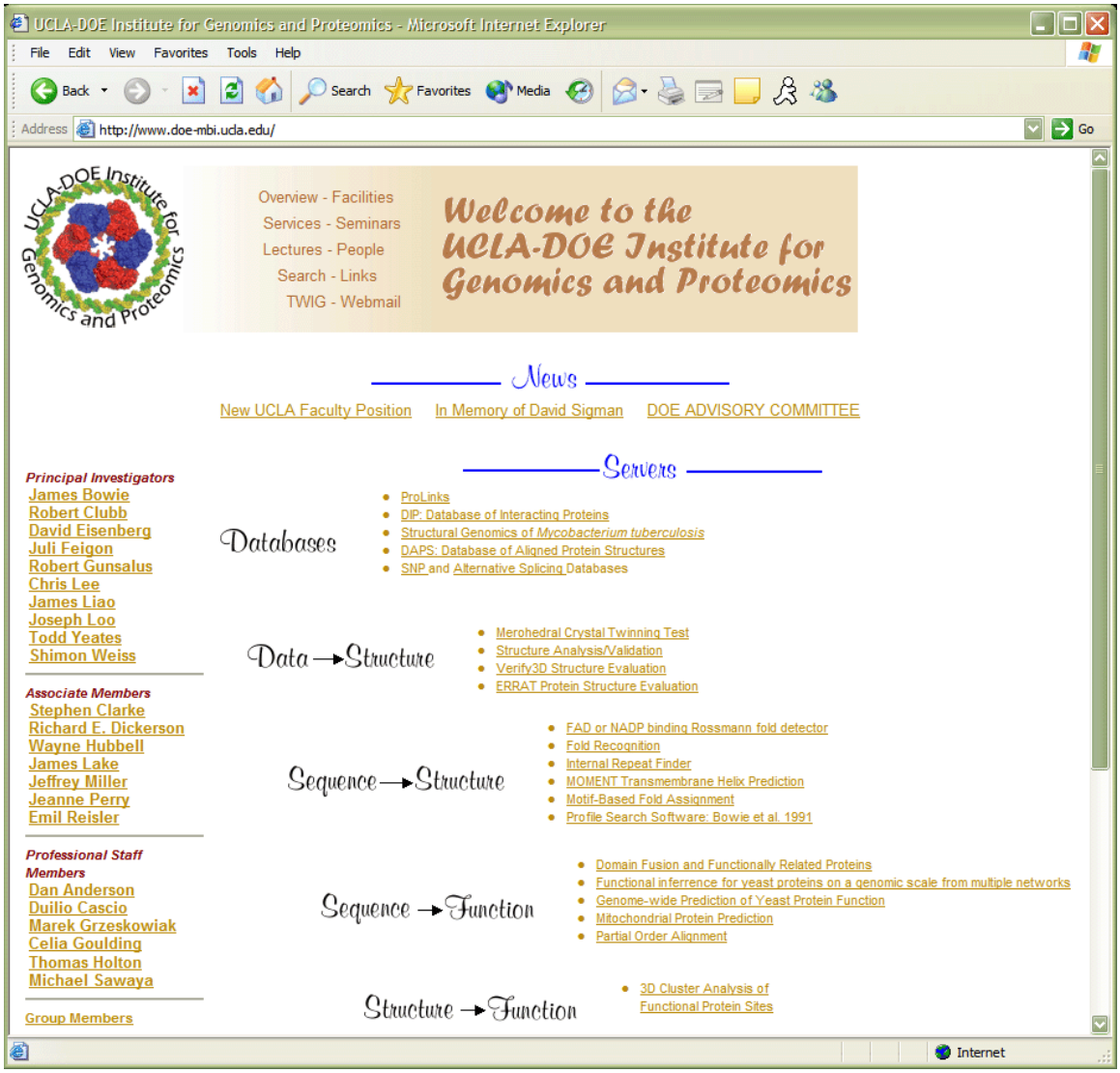
Center	PI Director	Manager
Computer-Bioinformatics	Christopher Lee	Dr. Duilio Cascio
Protein Expression	Jeanne Perry	Dr. Jeanne Perry
X-ray	Todd Yeates	Dr. Duilio Cascio
NMR	Robert Clubb, Juli Feigon	Dr. Robert Peterson
Biochemistry Instrumentation	Joseph Loo, Emil Reisler	Martin Phillips
Optical studies (just starting)	Shimon Weiss	Laurent Bentolila

All Centers sponsored and run by DOE but supported by multiple sources
Other Campus users welcomed; no user fees except computer service

V. The Web Site of the UCLA-DOE Institute: http://www.doe-mbi.ucla.edu

The Institute strives to make its numerous advances in technology open to the scientific community, and in particular, to other DOE units. The main mechanism for this is our extensive Web Site, the home page of which is shown in Fig. 4.

Fig. 4 The UCLA-DOE home page.



Publications:

"A Combined Algorithm for Genome-wide Prediction of Protein Function" Marcotte, E.M., Pellegrini, M., Thompson, M.J., Yeates, T.O., and Eisenberg, D., *Nature*, 402, 83-86 (1999).

"Protein Function in the Post-Genomic Era" Eisenberg, D., Marcotte, E.M., Xenarios, I., and Yeates, T. *Nature*, 405, 823-826 (2000).

"Genomic Evidence that the Intracellular Proteins of Archael Microbes Contain Disulfide Bonds." Mallick, P., Boutz, D. R., Eisenberg, D., and Yeates, T.O., *PNAS*, 99, 9679-9684 (2002).

“Protein Interactions: Two Methods for assignment of the reliability of high-throughput observations”, Deane, C.M., Salwinski, L., Xenarios, I., Eisenberg, D., *Molecular and Cellular Proteomics*, 1.5, 349-356 (2002).

"Describing biological protein interactions in terms of protein states and state transitions: the Live DIP database" Duan, X.J., Xenarios, I., and Eisenberg, D., *Molecular and Cellular Proteomics*, 1, 104-116 (2002).

"Helix-bundle membrane protein fold templates" James U. Bowie, Protein Science, 8, 2711-2719, (1999).

"Helix packing in Membrane Proteins" James U. Bowie, J. Mol. Biol. 272, 780-789, (1997).

Connolly KM, Ilangovan U, Wojciak JM, Iwahara M and **Clubb RT**. Major Groove Recognition by Three-Stranded β-Sheets: Affinity Determinants and Conserved Structural Features.*Journal of Molecular Biology, 300* 2000; 841-856.

Wojciak JM, Iwahara J and **Clubb RT**. The Solution Structure of the Mu repressor-DNA complex contains an immobilized 'wing' within the minor groove. *Nature Structural Biology, 8:* 84-90, (2001).

Ilangovan U, Iwahara J, Ton-That H, Schneewind O and **Clubb RT.** Assignment of the ^{1}H, ^{13}C and ^{15}N Signals of Sortase. *Journal of Biomolecular NMR, 19:*379-380, (2001).

Iwahara J Wojciak JM, and **Clubb RT.** Improved NMR spectra of a protein-DNA complex through rational mutagenesis and the application of a sensitivity optimized isotope-filtered NOESY experiment. *Journal of Biomolecular NMR, 19* 2001; 231-241.

Ilangovan U, Iwahara J, Ton-That H, Schneewind O and **Clubb RT.** Structure of Sortase, the transpeptidase that anchors proteins to the cell wall of *Staphylococcus aureus*. *Proceedings of the National Academy of Sciences (USA) 98* 2001; 6056-6061.

"Genome-wide detection of alternative splicing in expressed sequences of human genes", Barmak Modrek, Alissa Resch, Catherine Grasso and Chistopher Lee, Nucleic Acids Research, 29, (2001).

"Genome-wide analysis of single-nucleotide polymorphisms in human expressed sequences" Kris Irizarry, Vlad Kustanovich, Cheng Li, Nik Brown, Stanley Nelson, Wing Wong & Chistopher J. Lee, Nature America, 26, (2000).

"Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene affects", George C. Tseng, Min-Kyu Oh, Lars Rohlin, James C. Liao, and Wing Hung Wong, Nucleic Acids Research, 29, 2549-2557, (2001).

"Electrospray ionization mass spectrometry: a technology for studying noncovalent macromolecular complexes", Joseph A. Loo, International Journal of Mass Spectrometry, 200, 175-186, (2000).

"Nanohedra: Using symmetry to design self assembling protein cages, layers, crystals, and filaments" Jennifer E. Padilla, Christos Colovos, and Todd O. Yeates, PNAS, 98, 2217-2221, (2001).

"The Crystal Structure of Adenylosuccinate Lyase from Pyrobaculum aerophilum Reveals an Intracellular Protein with Three Disulfide Bonds", Eric A. Toth, Carolyn Worby, Jack E. Dixon, Eric R. Goedken, Susan Marqusee, and Todd O. Yeates, J. Mol. Biol, 301, 433-450, (2000).

VI. Mutual Cooperation of the UCLA-DOE Institute for Genomics and Proteomics (IGP) and DOE OBER

http://www.doe-mbi.ucla.edu/

The UCLA-DOE Institute receives the opportunity to participate in the scientific mission of DOE in the area of genomics and proteomics. This permits us to develop new computational and experimental technologies to advance the DOE mission, and to apply the new technologies to organisms and systems of interest to DOE.

Because of a long-standing agreement between the Institute and UCLA, the Cooperative Agreement between DOE and the Institute provides both direct and some indirect funding for research. This funding supports the 5 Core Technology Centers, which benefit Institute research and the entire UCLA Campus.

In recent years, the UCLA-DOE Institute has made several notable contributions to DOE missions. These include:

- Genome-wide 3D protein structure prediction*: a part of DOE's mission that got its start at UCLA
- Genome-wide inference of protein interaction networks*: a part of DOE's mission that got its start at UCLA
- Development of Partial-Order Alignment and other algorithms
- Large scale protein expression: pioneered in part by the UCLA-DOE Protein Expression Lab
- Web-accessible databases:
  Database of Interacting Proteins
  Database of Distantly Aligned Proteins
  SNP and Alternative Splicing Database
- Web-accessible servers: http://www.doe-mbi.ucla.edu/
  For Data->Structure, Sequence->Structure, Sequence->Function, and 3D Structure->function

*These UCLA-DOE methods are SCALABLE and can be applied to any DOE fully sequenced genome.