Project no. 517991
Project Acronym COSBIOM
Project Title: Towards Excellence in Computational Structural Biology and Biomaterials

Specific Support Action (SSA)
Life Sciences, Genomics and Biotechnology for Health

Publishable Final Activity Report

Period covered from 01.05.2005 to 30.04.2009 Date of Preparation: 12.06.2009

Start date of the Project: 01.05.2005 Duration: 4 years

Project Coordinator
Prof. Türkan Haliloğlu
Polymer Research Center

Revision [1]

1. Project execution

Contractor:

The Polymer Research Center (PRC) at Bogazici University (BU) in Istanbul, Turkey was established in 1989. The major research activities of PRC comprise of modeling and simulations of biological and synthetic macromolecules, with emphasis on biomolecular structure and dynamics, and the synthesis and characterization of multifunctional polymers with bio-applications. PRC has gained national and international recognition through the scientific accomplishments of its members, who have received numerous national awards and international recognition. The interdisciplinary research activities are conducted in the laboratories of PRC and other integrated laboratories, with joint participation of the Engineering and Arts and Sciences Faculty at BU. PRC forms an infrastructure for bringing together several distinguished researchers with diverse backgrounds in macromolecular physics/chemistry, molecular modeling/simulation/design, and polymer synthesis/characterization.

When the calls for the FP6-ACC-2004-SSA-2 were made, PRC submitted the proposal "Towards Excellence in **Co**mputational **S**tructural **Bio**logy and **Biom**aterials" (acronym "**COSBIOM**"). PRC received a successful evaluation from the call and was recognized as one of the eleven centers of excellence in Turkey. An EU Sixth Framework Programme Specific Support Action project contract (contract no: 517991) was duly signed by the INCO Research Directorate and the vice president of BU in April 2005. The project had originally a budget of the 849,960 Euros covering originally the period of May 2005 until April 2008. Extention request of the project duration from 36 months to 48 months was accepted by the EC Commission in February 2008. The major reason for the extension request was PRC's inability to undertake one of the work packages in the contract, WP7 (Molecular Modeling Workshop), which was scheduled to be held in September 2007. Within the scope of the first objective, the organization of an international workshop in molecular modeling was planned that would familiarize researchers in life sciences with the modern multi-scale computational/simulation techniques. Due to our intense efforts towards progress on other workpackages, the execution of this workshop had been postponed to the end of April 2009. Although the program of the workshop was being finalized by contacting with potential lecturers, following P3, due to the uncertainities regarding the hold up of the rest of the payment by EU, we couldn't commit to the lecturers for the Molecular Modeling Workshop (WP7) to be held in April 2009. The financial issue could only be clarified as of March 9, 2009, which left very limited time for the execution of this activity and we had to cancel it. The scientific advisor were informed regarding this issue.

Coordinator Contact Details:

Prof. Turkan Haliloglu
Address : Polymer Research Center,
Bogazici University, Bebek,
34342, Istanbul, Turkey
Phone : (+90) 212 359 70 02
Fax : (+90) 212 257 50 32
PRC website : http://www.prc.boun.edu.tr
COSBIOM website : http://www.prc.boun.edu.tr/cosbiom

1.1. Project objectives

The main objectives of the project COSBIOM are (1) to enhance PRC's excellence in the fields of computational structural biology and biomaterials by networking with European Research Centers through exchange of researchers (visiting scientists, postdoctoral fellows); (2) to increase PRC's research capacity by upgrading existing computational and experimental resources and by hiring postdoctoral fellows and graduate students; (3) to ensure the continuity of PRC's overall capacity within the scope of this support action by joint research activities and preparing new project proposals with collaborators in Europe; and (4) to disseminate all research results through scientific meetings, publications in peer-reviewed journals, and continuous posting of activities and results at website, and by providing novel computational methodologies to the scientific world via the grid infrastructure to be implemented on local computer network.

1.2 Work performed and major results

- ✓ The existing computational resources were upgraded through the purchase of DELL Poweredge web/mail and backup servers, tape library, HP MSA Storage and necessary software. Also purchase and installation processes of a new SGI ALTIX 350 shared memory parallel computer, a DELL Poweredge grid server and seven DELL Blade servers were completed. Additionally, during P4 with the extention of the project and the possible use of the budget remaining from some of the other Workpackages, the computational resources were further upgraded with the purchase and implementation of a parallel computing cluster composed of HP BL460c G6 Blade servers using new core i7 based Intel Xeon 5500 series processors, a file server, an MSA storage unit and components for the infrastructure including infiniband components and HP BladeSystem c7000 Enclosures.

- ✓ A gel permeation chromatography (GPC) was purchased to upgrade our research labs for synthesis of new biomaterials and analyze polymer samples. A standard operation procedure has been developed for new users. The operation procedure was also disseminated to two other research groups in BU other than the researchers in the COSBIOM consortium. This instrument has now been extensively utilized by several researchers.

- ✓ PRC's human potential has been increased through hiring of one post-doctoral fellow in M05, two post-doctoral fellows in M34 and one post-doctoral fellow in M37. While approaching to the end of the project, the latter post-doctoral fellow was awarded with a fellowship from

TUBITAK in March 2009 with a favourable proposal evaluation, having the opportunity to continue her research activities at PRC. Another postdoctoral fellow started at PRC with a similar fellowship awarded by TUBITAK. Activities of the post-doctoral fellows consist of computational biology projects: (1) The conformational properties and the role of the collective structural motions and the pH dependency on reactivity of dehydrofolate reductase, which is the target of several antitumor and antimicrobial drugs; (2) The study of GTP binding protein EF-TU by combined computational tools such as Molecular Dynamic (MD) simulations, Quantum Mechanic/Molecular Mechanic (QM/MM) calculations and Normal Mode (NM) analysis; (3) The development of the mixed coarse-graining approach recently introduced to the elastic network models, the investigation of the structure-function relationship of very large biological assemblages such as the ribosome at atomistic level with this computationally efficient model; (4) The dynamic behavior of binding sites and dynamics across the interface of protein complex structures, the identification of different types of interfaces; crystal interface, obligatory and transient interfaces; (5) The incorporation of the protein flexibility on the docking studies using elastic network model, where distinct conformers of protein Cyclophilin A is generated by the collective normal modes,whereas the ligand is held flexible at their side chains, the application of the same approach to the enzyme triosephosphate isomerase; (6) The nonlinear dynamics of apoptosis.

- ✓ PRC had over 20 visitor scientists in the 4 years duration of the project. The scientists of PRC has hosted in the first two years are Prof. Lennart Nilsson from the Molecular Modeling Group at Karolinska Institute, Prof. Martyn Ford, director of the Center for Molecular Design University of Portsmouth, Hampshire, Prof. Norbert Moszner from Ivoclar Vivadent AG, Liechtenstein, Prof. Mark Vincent from University of Manchester, Prof. Tim Clark from University of Erlangen, Prof. Walter Thiel from Max-Planck-Institute, Prof. Nir Ben Tal and Prof. Ruth Nussinov from Tel Aviv University, Prof. Veronique Van Speybroeck from University of Ghent, and Dr. Arlette Delay and Pascal Mieville from University of Geneva. In the third reporting period, PRC has hosted Prof. Nir Ben-Tal and Prof. Ruth Nussinov from Tel Aviv University, Prof. Gerald Monard from Université Henri Poincare, Prof. Dario Pasini from University of Pavia, and Prof Ulrike Salzner from Bilkent University. In the final reporting period, Dr. Paul Bates from Cancer Research UK, Dr. Ewald Pauwels from University of Gent, Prof. Alexandre Bonvin from Utrecht University and Prof. Assa Lifshitz from Hebrew University has visited PRC. Also, three scientists attending the national conference HIBIT'08, Prof. Alfonso Valencia from Spanish National Cancer Research Centre (CNIO), Prof. Rita Casadio from University of Bologna and Prof. Ruth Nussinov from Tel Aviv University, were supported by COSBIOM Project.

- ✓ There is an on-going collaboration on modeling the kinetics of Free Radical Polymerization (FRP) between PRC and Center of Molecular Modeling at Ghent University. The collaboration between PRC and the Molecular Modeling Group at Karolinska Institute is focused on the analysis of structure-function relationships of restriction endonucleases. Center for Molecular Design from University of Portsmouth is focused on computer-aided drug design, QSAR, bioinformatics, application, formulation, pharmacokinetics, mode of action and ecotoxicology of crop protectants, harvesting energy for use with biological systems. Unfortunately, due to Prof. Ford's decease, the future projects that will be undertaken during the next years are delayed. We initiated collaboration with Prof. Norbert Moszner, Ivoclar Vivadent AG, Liechtenstein, in the field of new strongly acidic, free-radically polymerizable monomers, such as phosphonic acids for dental applications. A proposal describing the structure of the monomers was prepared by Prof. Avci. After approval of the proposal by Prof. Moszner a development agreement was signed between both parties. During the visit of

Prof. Mark Vincent to PRC, both parties have discussed their common project on the behavior of the Evans chiral bis(oxazoline)copper(II) ([Cu_{II}box]) catalyst in relation to the Claisen rearrangement. Prof. Walter Thiel has discussed the usage of the QM/MM methodology with PRC group during his visit. This visit has enabled PRC group to learn the capabilities and frontiers of the QM/MM methodologies. Visits between Prof. Nir Ben-Tal's group and PRC have further enhanced and also initiated the new research activities between the two groups. Combining novel computational tools developed in these two laboratories in the study of protein interfaces, the formulation of the network of interactions, and the prediction of the shape of transmembrane helices, summarizes the new research topics on this collaboration. With Prof. Ruth Nussinov's visits, the project on protein interaction networks is progressing. Prof. Can Ozturan, who is also a member of COSBIOM project team, and one of his PhD students and two master students have also been involved in this project. The visit of Prof. Alexandre Bonvin has initiated the collaboration on the implementation of Gaussian network model into the docking algorithm of Prof. Bonvin's group. Prof. Gerald Monard (with Prof. Andre Dedieu and Manuel Ruiz Lopez) from Université de Henri Poincaré-Nancy has visited PRC, and a future collaboration on QM/MM studies with Prof. Monard's group is projected. Prof. Dario Pasini has been invited to give a seminar on copolymerization techniques for the synthesis of functional materials. After the visit of Dr. Ulrike Salzner as an invited speaker, future projects on modeling dental adhesive materials are discussed. Visit of Prof. Turkan Haliloglu to King's College London initiated a new collaboration with Prof. France Fraternali on the hub proteins in protein-protein interaction networks. A possibility of preparing a joint proposal was discussed.

- ✓ Set-up of web-based COSBIOM Project Information System has eased the information sharing between both COSBIOM project members and the interested audience. For this purpose, a content management system and some other third party components have been installed on a web server. Those who visit the web-site can see general information about the Sixth Framework Programme, PRC and COSBIOM project and its workpackages, the project consortium and the academic personnel, and technical advisory board. Upcoming events like seminars, meetings are announced in the website. Also past events can be viewed together with documentations which may be useful for those who are interested in the topics. The posters that are presented by the PRC staff in various conferences are also given in documentations section.

- ✓ PRC disseminates knowledge through attending conferences in Europe. In the four years duration of the project, many graduate students were supported to participate in conferences.
  - Two PhD students participated in the Fifth European Workshop in Drug Design, which was held in Certosa di Pontignano, Siena, Italy with poster presentations. The presentations were on mixed coarse-grained elastic network model analysis.
  - Prof. Viktorya Aviyente, a core member of the COSBIOM consortium, participated with four of her graduate students in the 11th International Conference on the Applications of Density Functional Theory in Chemistry and Physics in Geneva, Switzerland. The studies presented in this conference are on the structure-reactivity relationship of acrylic monomers investigated with the B3LYP/6-31G* methodology.
  - A PhD student attended the SFC-Eurochem conference held in Nancy in M09 and presented a poster entitled "The Elucidation of the Deamidation Mechanism in Peptides".
  - In the second year of the project, nine graduate students from PRC were supported to participate in conferences in Europe (supported by COSBIOM Project) and 12 graduate

students participated in meetings held elsewhere- national or in the U.S.A. (supported by other sources).

- In the third year of the project, one post-doctoral fellow and six graduate students were supported to participate in meetings held in Europe, and seven graduate students participated in meetings held in USA (supported by other sources).
- A graduate student from PRC had a visit to the Manchester University Chemistry Department in M08 in the scope of a collaborative project on a novel catalyst [Cu(II)box].
- Aarhus University Department of Chemistry in Denmark invited one of PRC's graduate students in order to work on a collaborative research project concerning the enzymatic reaction mechanism of benzoylformate decarboxylase by molecular dynamics methods.
- A PhD student had a visit to the Theoretical Chemistry Group in the University of Henri Poincare, Nancy.
- A PhD student had a visit to the laboratory of Prof. Van Speybroeck in the University of Ghent in M29.
- In the fourth year of the project, Prof. Viktorya Aviyente made oral presentations in XII National Chemistry Meeting, North Cyprus and WATOC08, Sydney, Australia (supported by other sources).
- In P4, Eight graduate students has presented posters in conferences and workshops like European Conference on Computational Biology (ECCB), Cagliari, Sardinia-Italy; 2nd Euchems Chemistry Congress, Torino, Italy; 6th Congress on Electronic Structure: Principles and Applications - ESPA 2008, Palma, Spain; and Biomolecular Simulation, EMBO Practical Course, Pasteur Institute, Paris, France.

✓ Dissemination of knowledge to the scientific community has been through publications in peer-reviewed journals. Citations to these publications show how successfully the knowledge has been disseminated. In the four years duration of the project, 75 papers were published. In the same period, papers of PRC and the members of COSBIOM had 1251 citations.

✓ Grid infrastructures, another tool for dissemination of the knowledge, enable various resources such as computers, storage systems, data resources, software and special devices to be shared in what is called virtual organizations. Computational biology, biomaterials, chemistry, bioinformatics applications make sample use of such resources and hence the development of a grid for sharing these resources is an appropriate step to take, which will strengthen the relationships among scientists in these areas. Installation of the grid software has been completed and is now fully functioning.

1.3 Impact of the project on its industry or research sector

The foreseen benefits of the project are immediate increase (1) in the competency level in developing new computational methodologies in the fields of structural biology and bioinformatics in collaboration with experimental and computational groups in Europe for the solution of important health-related problems, (2) in collaborations related to the design of stereo-selective synthesis of drugs, and design/synthesis of novel multifunctional polymers with applications in dentistry and nanotechnology, and (3) in computational, experimental and human resources, which will strengthen PRC's research infrastructure and its potential for future participation in the European research platform.

The current project is relevant to several thematic priorities of the Sixth Framework Programme, in particular advanced genomics and its applications for health, bioinformatics, basic biological processes, knowledge-based multifunctional materials, and biomaterials. Through the COSBIOM project and potential participation in the Seventh Framework Programme, PRC will contribute to both European research and the research and technological development in Turkey.

2. Dissemination and use

Section 1. Exploitable knowledge and its use

Exploitable knowledge (description)	**Exploitable product(s) or measure(s)**	**Sector(s) of application**	**Timetable for Commercial Use**	**Patents or other IPR protection**	**Owner& Other partners(s) involved**
Prediction of mechanistically informative regions of protein structures	HingeProt-server	Scientific Community	2006	Published in a peer-reviewed journal*	BU-PRC
Grid	Grid	Scientific Community	2007		BU-PRC
Prediction of protein stability using machine learning technique	Prediction of protein stability server (Mlsta)	Scientific Community	2008	Under review in a peer-reviewed journal**	BU-PRC
A framework of a database for protein-protein interactions	Database of molecular Interactions (MIMCITY)	Scientific Community	2009	In preparation	BU-PRC

* Emekli U, Schneidman-Duhovny D, Wolfson HJ, Nussinov R, Haliloglu T. (2008) HingeProt: Automated Prediction of Hinges in Protein Structures. *Proteins* 70(4), 1219.
** Ozen, A, Gonen, M, Alpaydin, E, Haliloglu T. Machine Learning Integration for Predicting the Effect of Single Amino Acid Substitutions on Protein Stability. Under review in *Bioinformatics*

HingeProt:

http://www.prc.boun.edu.tr/appserv/prc/hingeprot/

HingeProt is a server developed in the previous reporting period for prediction of mechanistically informative regions of protein structures.

Proteins are highly flexible molecules. It is common to classify protein motions into shear and hinge motion. Shear motions are very limited and involve large number of residues. On the other hand, hinge motions are similar to rotations around an articulated joint and therefore can be very large. Hinge motion is characterized by large changes in main-chain torsional angles occurring at a localized region, which is called ***a hinge***. Hinge motions usually involve a small number of residues, since even one bond can provide the required rotational freedom. This kind of protein motion is free of packing constraints. When a chain exhibits hinge motion at the region connecting two structural domains, each domain behaves as a rigid body and packing interactions can appear/disappear between the interfaces of those rigid bodies.

Hinge motions usually occur upon binding to another molecule, or upon activation/deactivation of the protein.

Therefore, hinge regions are the mechanistically informative regions of the structure and are of great importance in mediating cooperative motions that have functional importance.

HingeProt is a web server for predicting rigid protein parts and the flexible hinge regions connecting them in the native topology of protein chains by employing elastic network (EN) models. *HingeProt* makes use of both Gaussian Network Model (GNM) and Anisotropic Network models (ANM).

HingeProt server focuses on the prediction of the rigid parts and the hinge regions using a **single** static conformation of a protein structure. The hinge regions are the mechanistically informative regions of the structure and are of importance in mediating cooperative motions that have functional importance. GNM calculates the mean-square fluctuations and the correlation between the fluctuations of residues in the most dominant (slowest two) modes, which were shown to overlap with known protein motions. These suggest hinge regions and the cooperation between them. ANM provides the direction of the fluctuations in the corresponding modes.

HingeProt is expected to be useful in a range of potential applications, especially in prediction protein-protein association, flexible docking and in refinement of the structure of the modeled complexes. *HingeProt* predictions are also helpful in fitting flexible hinge-bent protein structures into EM density maps and refining the EM structures. In addition, hinge regions can help in understanding functional mechanisms of macromolecular structures and assemblies.

Given an input protein chain, *HingeProt* identifies the rigid parts and the hinges connecting them, and the direction of the fluctuation of each residue in the slowest two modes.

PRC has developed the algorithm and the server in collaboration with the Tel-Aviv University Structural Bioinformatics group, Israel. There are continuous efforts for the development of the server. *Hingeprot* has now been extensively tested and the paper that describes the theoretical background and its use is just published. The updates in this reporting period are: improvement of server performance by achieving submitted structures for future usage and updating the visualization programs to prevent unexpected failures during runs; design of the server more user friendly by changing page layout and allowing the user to change some parameters.

The server is available for free public access. It does not have any economic market considerations (not a commercial product) and does not have a socio-economic impact as an exploitable result. *HingeProt* is a milestone for PRC in developing servers. Other servers will follow, arising further collaborations. Other researchers from scientific community may use this server in their studies, like docking etc. Around 1000 usage entries of *HingeProt* have been recorded for the period of July 2008-June 2009. As an example; *Hingeprot* was utilized in several studies. Some of the citations to these papers are given below:

Hu Pan, Joseph D. Ho, Robert M. Stroud, and Janet Finer-Moore, "The Crystal Structure of *E. coli* rRNA Pseudouridine Synthase RluE", Journal of Molecular Biology, 2007 367(5), 1459–1470.

Rakhi Agarwal, Stephen K. Burley and Subramanyam Swaminathan,"Structural Analysis of a Ternary Complex of Allantoate Amidohydrolase from Escherichia coli Reveals its Mechanics", Journal of Molecular Biology, 2007, 368, 450-463.

Vajda S, Kozakov D, "Convergence and combination of methods in protein-protein docking", CURRENT OPINION IN STRUCTURAL BIOLOGY, 19:2, 164-170, APR 2009.

Skjaerven L, Hollup SM, Reuter N, "Normal mode analysis for proteins", JOURNAL OF MOLECULAR STRUCTURE-THEOCHEM, 898:1-3 Sp. Iss., 42-48, MAR 30 2009.

Rodziewicz-Motowidlo S, Iwaszkiewicz J, Sosnowska R, et al., "The Role of the Val57 Amino-Acid Residue in the Hinge Loop of the Human Cystatin C. Conformational Studies of the Beta2-L1-Beta3 Segments of Wild-Type Human Cystatin C and Its Mutants", BIOPOLYMERS, 91:5, 373-383, MAY 2009.

Jimenez A, Clapes P, Crehuet R, "Protein Flexibility and Metal Coordination Changes in DHAP-Dependent Aldolases", CHEMISTRY-A EUROPEAN JOURNAL, 15:6, 1422-1428, 2009.

Andrusier N, Mashiach E, Nussinov R, et al., "Principles of flexible protein-protein docking", PROTEINS-STRUCTURE FUNCTION AND BIOINFORMATICS, 73:2, 271-289, NOV 1 2008.

MIMCITY:

http://erciyes.cmpe.boun.edu.tr:8080/MIMCITY/

MIMCITY is a platform (in preparation) that offers private as well as shared database of molecular interactions maps. Users can create accounts in MIMCITY, enter molecular interaction maps and share them with other users. Search mechanisms will also be provided to locate relevant interaction maps of other users. The interface to MIMCITY will be provided also by various web services. Provision of a web services-based interaction with the database allows the users to actually write automated programs that can access and search our database. This is especially useful when repeated interactions with the database need to be done.

All the data passed between the MIMCITY database server and the clients are stored in XML format. MIMCITY has been implemented using the Java language and various open-source tools.

GRID:

The share of our computational, software and data resources with our collaborators and other external users is achieved by a hybrid computational and data grid infrastructure called COSBIOM that has been installed as of February 2008. Existing computational applications are now grid-enabled, as well as our newly acquired computational servers and data are available as resources on our grid. The COSBIOM grid not only enables our resources to be

shared, but also enables us to network and integrate more easily with other Biology grids in Europe.

PRC has completed the establishment of the grid infrastructure with its own human resources. The grid has no economic considerations since it has not been established as a commercial product. Its establishment aims to easily collaborate with other research centers throughout Europe and to disseminate knowledge.

Mlsta: Machine Learning Integration for Predicting the Effect of Single Amino Acid Substitutions on Protein Stability:

Computational prediction of protein stability change due to single-site amino acid substitutions is of interest in protein design and analysis. The performance of the currently available predictors can be improved in two ways; (1) considering additional sequence and structure based features, (2) implementing different machine learning integration approaches to combine information from different features or representations. The amino acid substitution likelihoods, the equilibrium fluctuations of the alpha- and beta-carbon atoms, and the packing density are the features that we incorporate. Three different approaches are investigated: **early**, **intermediate**, and **late integration**, which respectively combine features, kernels over feature subsets, and decisions.

In multikernel support vector machines, kernels can be combined by two different methods: We can calculate kernel functions on different representations or calculate different kernel functions on the same representation. One can take a sum over different kernels and summation rule is applied successfully in computational biology where heterogeneous data sets exist by the nature of the biological problems. Lanckriet *et al.* (Journal of Machine Learning Research, 5, 27-72, 2004) and Bach *et al.* (ICML 2004 Proceedings) propose to replace kernel function with a weighted summation of kernel functions and the combination weights are new parameters optimized in training. In addition to the flexibility of constructing weighted combination rules, using multikernel SVMs provides two important advantages: (1) Information can be extracted about the classification task at hand. The feature sets used in kernel functions with larger weights give more relevant information in terms of classification. For example, obtaining information about important features in biological problems such as disease diagnosis and drug development is as important as classification accuracy. (2) Kernel functions with zero weights can be eliminated. If such feature sets are obtained by using costly and time consuming experimental procedures, eliminating them is useful.

We use the mathematical model developed for multiple kernel learning by Bach *et al.* and implement it using MOSEK optimization software. We combine the Gaussian kernels evaluated on different feature subsets and their width parameters are validated from various multiples of the average nearest neighbor distances in the corresponding feature subsets.

This server will be available for public service upon publication of the related manuscript by Ozen *et al*, now it's under construction.
http://www.prc.boun.edu.tr/appserv/prc/mlsta).

Section 2. Dissemination of Knowledge

Planned/actual Dates	Type	Type of audience	Countries addressed	Size of audience	Partner responsible/ involved
M01	Publications	Research	All countries	N/A	PRC
April Issue	Press release – Article in Parliament Magazine	General public	All European countries	N/A	PRC
M11	Project website	General public	All countries	N/A	PRC
November 23rd, 2005	Seminar Dr. Veronique Von Speybroeck	Research	Turkey	25	PRC
March 13 - 17, 2006	Seminar Prof. Lennart Nilsson	Research	Turkey	25	PRC
April 24 - 27, 2006	Seminar Prof. Martyn Ford	Research	Turkey	25	PRC
May 17th, 2006	Seminer Prof. Norbert Moszner	Research	Turkey	25	PRC
May 30th, 2006	Seminar Prof. Mark Vincent	Research	Turkey	25	PRC
May 17th, 2006	Seminar Prof. Norbert Moszner	Research	Turkey	25	PRC
May 30, 2006	Seminar Dr. Mark Vincent	Research	Turkey	25	PRC
October 31st, 2006	Seminar Dr. Tim Clark	Research	Turkey	25	PRC
March 8th 2007	Arlette Delay, Pascal Mieville	Research	Turkey	25	PRC
March 9th, 2007	Seminar Prof. Walter Thiel	Research	Turkey	25	PRC
April 26th, 2007	Seminar Dr. Nir Ben Tal	Research	Turkey	25	PRC
April 26th, 2007	Dr. Veronique Van Speybroeck	Research	Turkey	25	PRC
Oct 4th, 2007	Seminar Prof. Ulrike Salzner	Research	Turkey	25	PRC
Nov 27th, 2007	Seminar Prof. Dario Pasini	Research	Turkey	25	PRC
December 11th, 2007	Seminar Prof. Gerald Monard	Research	Turkey	25	PRC
May 14th, 2008	Seminar Prof. Assa Lifshitz	Research	Turkey	25	PRC
May 22nd, 2008	Seminar Prof. Alexandre Bonvin	Research	Turkey	25	PRC
March 31st, 2009	Seminar Dr. Ewald Pauwels	Research	Turkey	20	PRC
April 28th, 2009	Seminar Dr. Paul Bates	Research	Turkey	30	PRC
M01	Posters	Research	Italy	N/A	PRC
M05	Posters	Research	Switzerland	N/A	PRC
3-8 September 2006	Posters	Research	Italy	N/A	PRC
18-21 May,2006	Posters	Research	Turkey	N/A	PRC
18-21 May,2006	Oral Presentation	Research	Turkey	N/A	PRC
10-13 July 2006	Poster	Research	France	N/A	PRC
23-30 March 2007	Oral Presentations	Research	Chicago,US	N/A	PRC
12-16 May 2007	Posters	Research	Sweden	N/A	PRC
19-30 June 2007	Oral Presentation	Research	Italy	N/A	PRC
1-5 July 2007	Posters	Research	Slovenia	N/A	PRC

26-30 August 2007	Posters	Research	The Netherlands	N/A	PRC
6-9 November 2007	Oral Presentation	Research	The Netherlands	N/A	PRC
24-25 April 2008	Posters	Research	Turkey	N/A	PRC
1-8 July 2008	Poster	Research	France	N/A	PRC
2-5 September 2008	Posters	Research	Spain	N/A	PRC
16-20 September 2008	Posters	Research	Italy	N/A	PRC
22-26 September 2008	Posters	Research	Italy	N/A	PRC
September 2008	Oral Presentation	Research	Australia	N/A	PRC
October 2008	Oral Presentation	Research	North Cyprus	N/A	PRC
April 2009	Oral Presentation	Research	Turkey	N/A	PRC

- ✓ References of journal publications are given in Deliverable 4.2.

- ✓ Dr. Veronique Von Speybroeck gave a seminar on "Chemical kinetics of organic and organometallic reactions through molecular modeling" on November 23rd, 2005.

- ✓ Prof. Lennart Nilsson; the director of the Molecular Modeling Group at Karolinska Institute, Sweden, gave a seminar on "Molecular Dynamics - Femtoseconds to Microseconds" on March 14th and a tutorial on the usage of macromolecular simulation program CHARMM.
- ✓ Prof. Martyn Ford; the director of the Center for Molecular Design, University of Portsmouth, gave a seminar on "Modelling Molecules Using Local Surface Properties and Motion" on April 25th 2006.
- ✓ Prof. Moszner gave a seminar entitled 'New Components for Dental Filling Materials' on May 17th, 2006.
- ✓ Dr.Mark Vincent gave a seminar on "The Wet and 'Dry' Fluoride ion: Physical Structure and Reactivity" on May 30, 2006.
- ✓ Dr.Clark gave a talk on "Modeling and cheminformatics using purely surface properties" on October 31st, 2006, and has introduced the usage of their new software in conjunction to predict properties based on surface properties.
- ✓ Dr. Thiel gave a talk on "Combined QM/MM studies of enzymes" on March 9, 2007. and has discussed the usage of the QM/MM methodology with Dr.Aviyente's group.
- ✓ Dr. Nir Ben Tal gave a talk on "Modeling the structure of the EmrE multidrug transporter: when two x-ray structures are not enough" on April 26th, 2007 at BU.
- ✓ Dr. Van Speybroeck has visited the Chemistry Department of Bogazici University and gave a talk on "Simulating reactions in organic solvents".
- ✓ Arlette Delay, Pascal Mieville, they gave a seminer on "The Safety, Handling and Waste Management of Biomaterials and Chemical Hazards" on March 8th 2007.
- ✓ Dr. Ulrike Salzner gave a talk on her results on modeling the doping processes in polythiophene and thiophene oligomers.
- ✓ Dr. Daroi Pasini gave a seminar on "Cyclopolymerization as a Tool for the Synthesis of Functional Materials".
- ✓ Dr. Paul Bates, who is the head of Biomolecular Modelling Laboratory at Cancer Research Institute UK London Research Institute, visited PRC on April 27-29, 2009. He gave a seminar titled 'Protein-Protein Interaction in their Social Context'. Possible future collaborations related to elastic network models and protein-protein docking have been discussed. Currently some of the elastic network model related algorithms

developed at PRC are being used in the Biomolecular Modelling Laboratory by a postdoctoral fellow, who is a former Ph.D. graduate of PRC.

- ✓ Dr. Ewald Pauwels gave a seminar on "Usage of hybrid computational tools to model spectroscopic properties of enzyme-substrate complexes".
- ✓ Prof. Alexandre Bonvin gave a seminar on "Information-driven Modeling of Biomolecular Complexes" and the protein docking program HADDOCK developed in Utrecht University.
- ✓ Prof. Assa Lifshitz gave a seminar on "Anatomy of Complex Reaction Systems. Combustion Reaction Mechanisms from Ignition Delay Times".
- ✓ Under the organization of the conference HIBIT'08,
    - Prof. Alfonso Valencia, from Spanish National Cancer Research Centre (CNIO), gave a talk titled "Computational Methods for the prediction of protein interactions at the genome scale, and for the extraction of interactions from published papers",
    - Prof. Rita Casadio, from University of Bologna, gave a talk titled "Protein Folding, Misfolding and Diseases: The I-Mutant Suite."
    - Prof. Ruth Nussinov, from Tel Aviv University, gave a talk titled "RNA structures: Prediction, Alignment and 3D Classified Database"

- ✓ Prof. Haliloglu collaborates with Prof. Ethem Alpaydin to develop novel algorithms to predict protein stability. A new web server will be available to the public use upon publishing the related manuscript by Ozen, A, Gonen M, Alpaydin E and Haliloglu T.

- ✓ Prof. Haliloglu collaborates with Prof. Ruth Nussinov and Dr. Can Özturan in the development of the platform that offers private as well as shared database of molecular interactions maps. It's a framework of a database for protein-protein interactions.

- ✓ Web-based COSBIOM Project Information System was set up during the first year of the project. The intention behind the system is to ease the information sharing between both COSBIOM project members and the interested audience. For this purpose, a content management system and some other third party components have been installed on a web server. The COSBIOM project information system is available at http://www.prc.boun.edu.tr/cosbiom/

  The system provides the following properties:
    - Creation, management, distribution, publishing, and discovery of information for the COSBIOM project.
    - Publishing news related to the COSBIOM project and announcements of the upcoming events.
    - Sharing documents and publications through a file repository.
    - Sharing internal information between COSBIOM project members.
    - Announcing and getting applications for open post doctoral positions.
    - Syndication

Section 3. Publishable Results

COSBIOM, in its first year, has generated a very useful exploitable result, **HingeProt**, available for all the researchers studying on bioinformatics and stuctural biology. HingeProt is a server developed for prediction of mechanistically informative regions of protein structures.

Proteins are highly flexible molecules. It is common to classify protein motions into shear and hinge motion. Shear motions are very limited and involve large number of residues. On the other hand, hinge motions are similar to rotations around an articulated joint and therefore can be very large. Hinge motion is characterized by large changes in main-chain torsional angles occurring at a localized region, which is called ***a hinge***. Hinge motions usually involve a small number of residues, since even one bond can provide the required rotational freedom. This kind of protein motion is free of packing constraints. When a chain exhibits hinge motion at the region connecting two structural domains, each domain behaves as a rigid body and packing interactions can appear/disappear between the interfaces of those rigid bodies. Hinge motions usually occur upon binding to another molecule, or upon activation/deactivation of the protein.

Therefore, hinge regions are the mechanistically informative regions of the structure and are of great importance in mediating cooperative motions that have functional importance.

HingeProt is a web server for predicting rigid protein parts and the flexible hinge regions connecting them in the native topology of protein chains by employing elastic network (EN) models. *HingeProt* makes use of both Gaussian Network Model (GNM) and Anisotropic Network models (ANM).

HingeProt server focuses on the prediction of the rigid parts and the hinge regions using a **single** static conformation of a protein structure. The hinge regions are the mechanistically informative regions of the structure and are of importance in mediating cooperative motions that have functional importance. GNM calculates the mean-square fluctuations and the correlation between the fluctuations of residues in the most dominant (slowest two) modes, which were shown to overlap with known protein motions. These suggest hinge regions and the cooperation between them. ANM provides the direction of the fluctuations in the corresponding modes.

HingeProt is expected to be useful in a range of potential applications, especially in prediction of protein-protein association by flexible docking and in refinement of the structure of the modeled complexes. *HingeProt* predictions are also helpful in fitting flexible hinge-bent protein structures into EM density maps and refining the EM structures. In addition, hinge regions can help in understanding functional mechanisms of macromolecular structures and assemblies.

Given an input protein chain, *HingeProt* identifies the rigid parts and the hinges connecting them, and the direction of the fluctuation of each residue in the slowest two modes.

HingeProt has been developed in collaboration with the structural bioinformatics group in Tel-Aviv University, Israel through information exchange. PRC shared its knowledge on elastic network models and got support from Tel-Aviv University, which is a very experienced group in developing servers in bioinformatics fields.

The server has been published on world-wide-web with the address; http://www.prc.boun.edu.tr/appserv/prc/HingeProt2/. It was also introduced to the scientific world with a paper published in a peer-reviewed journal.

In the third year of COSBIOM, another web based server for the prediction of the effect of single amino acid substitutions on protein stability using machine learning techniques, has been developed and currently is under construction (http:// http://www.prc.boun.edu.tr/appserv/prc/mlsta).

Computational prediction of protein stability change due to single-site amino acid substitutions is of interest in protein design and analysis. The performance of the currently available predictors can be improved in two ways; (1) considering additional sequence and structure based features, (2) implementing different machine learning integration approaches to combine information from different features or representations. The amino acid substitution likelihoods, the equilibrium fluctuations of the alpha- and beta-carbon atoms, and the packing density are the features that we incorporate. Three different approaches are investigated: early**,** intermediate**,** and late integration**,** which respectively combine features, kernels over feature subsets, and decisions.

In multikernel support vector machines, kernels can be combined by two different methods: We can calculate kernel functions on different representations or calculate different kernel functions on the same representation. One can take a sum over different kernels and summation rule is applied successfully in computational biology where heterogeneous data sets exist by the nature of the biological problems. Lanckriet *et al.* (Journal of Machine Learning Research, 5, 27-72, 2004) and Bach *et al.* (ICML 2004 Proceedings) propose to replace kernel function with a weighted summation of kernel functions and the combination weights are new parameters optimized in training. In addition to the flexibility of constructing weighted combination rules, using multikernel SVMs provides two important advantages: (1) Information can be extracted about the classification task at hand. The feature sets used in kernel functions with larger weights give more relevant information in terms of classification. For example, obtaining information about important features in biological problems such as disease diagnosis and drug development is as important as classification accuracy. (2) Kernel functions with zero weights can be eliminated. If such feature sets are obtained by using costly and time consuming experimental procedures, eliminating them is useful.

We use the mathematical model developed for multiple kernel learning by Bach *et al.* and implement it using MOSEK optimization software. We combine the Gaussian kernels evaluated on different feature subsets and their width parameters are validated from multiples of the average nearest neighbor distances in the corresponding feature subsets.

In the fourth year of the project, a platform named MIMCITY is in preparation. It's a platform that offers private as well as shared database of molecular interactions maps. The interface to MIMCITY will be provided by various web services. The web address of the site under construction is: http://erciyes.cmpe.boun.edu.tr:8080/MIMCITY/

Contact details:

Name:	Turkan Haliloglu	**Phone:**	90 212 359 70 02
Role:	Director	**Fax:**	90 212 257 50 32
Organization:	Bogazici University	**E-Mail:**	halilogt@boun.edu.tr
Department:	Polymer Research Center	**Website:**	http://prc.boun.edu.tr/cosbiom
Address:	Cevdetpaşa Caddesi No: 115, Bebek		
Postcode and City:	34342 Istanbul		