Final Progress Report for Research Projects Funded by Health Research Grants

Instructions: Please complete all of the items as instructed. Do not delete instructions. Do not leave any items blank; responses must be provided for all items. If your response to an item is "None", please specify "None" as your response. "Not applicable" is not an acceptable response for any of the items. There is no limit to the length of your response to any question. Responses should be single-spaced, no smaller than 12-point type. The report **must be completed using MS Word**. Submitted reports must be Word documents; they should not be converted to pdf format. Questions? Contact Health Research Program staff at 717-783-2548.

1. **Grantee Institution:** MPC Corporation

2. **Reporting Period (start and end date of grant award period):** 1/1/09-6/30/10

3. **Grant Contact Person (First Name, M.I., Last Name, Degrees):** Shannon M. Barnes, MS

4. **Grant Contact Person's Telephone Number:** 412-648-9676

5. **Grant SAP Number:** 4100047641

6. **Project Number and Title of Research Project:** Project 4, Modeling the Interactions of Dual Specificity Phosphatases with Inhibitors

7. **Start and End Date of Research Project:** 1/1/09-6/30/10

8. **Name of Principal Investigator for the Research Project:** Ivet Bahar, PhD

9. **Research Project Expenses**.

   9(A) Please provide the amount of health research grant funds spent on this project for the entire duration of the grant, including any interest earned that was spent:

   $ 28,932

   9(B) Provide the last names (include first initial if multiple individuals with the same last name are listed) of **all** persons who worked on this research project and were supported with health research funds. Include position titles (Principal Investigator, Graduate Assistant, Post-doctoral Fellow, etc.), percent of effort on project and total health research funds expended for the position. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

Last Name	Position Title	% of Effort on Project	Cost
Bakan	Graduate Student Researcher	16.67%, Yr1	$6,163.73
Bakan	Graduate Student Researcher	16.67%, Yr 2	$6,163.73
Dutta	Graduate Student Researcher	50%, Yr 2	$18,487.50

9(C) Provide the names of **all** persons who worked on this research project, but who *were not* supported with health research funds. Include position titles (Research Assistant, Administrative Assistant, etc.) and percent of effort on project. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

Last Name	Position Title	% of Effort on Project
Bahar	Principal Investigator	5

9(D) Provide a list of **all** scientific equipment purchased as part of this research grant, a short description of the value (benefit) derived by the institution from this equipment, and the cost of the equipment.

Type of Scientific Equipment	Value Derived	Cost

10. Co-funding of Research Project during Health Research Grant Award Period. Did this research project receive funding from any other source during the project period when it was supported by the health research grant?

Yes____X____ No_____

If yes, please indicate the source and amount of other funds:

"*In silico* screening strategies for identifying small molecule inhibitors for PUMA-mediated apoptosis" is currently funded through NIH (PI: Greenberger) as a pilot study for one year at $67,975. The grant is also based on drug discovery, and the methods developed in this project have been used in the development of p53 upregulated modulator of apoptosis (PUMA) inhibitors. This grant is funded through August 2010, and we have received additional funding to continue through 2016, as noted below.

11. Leveraging of Additional Funds

11(A) As a result of the health research funds provided for this research project, were you able to apply for and/or obtain funding from other sources to continue or expand the research?

Yes__X_____ No_____

If yes, please list the applications submitted (column A), the funding agency (National Institutes of Health—NIH, or other source in column B), the month and year when the application was submitted (column C), and the amount of funds requested (column D). If you have received a notice that the grant will be funded, please indicate the amount of funds to be awarded (column E). If the grant was not funded, insert "not funded" in column E.

Do not include funding from your own institution or from CURE (tobacco settlement funds). Do not include grants submitted prior to the start date of the grant as shown in Question 2. If you list grants submitted within 1-6 months of the start date of this grant, add a statement below the table indicating how the data/results from this project were used to secure that grant.

A. Title of research project on grant application	B. Funding agency (check those that apply)	C. Month and Year Submitted	D. Amount of funds requested:	E. Amount of funds to be awarded:
Mitochondrial targeting against radiation damage (CMCR renewal)	☑NIH ☐ Other federal (specify:_____ _____) ☐ Nonfederal source (specify: _____)	11/09	$898,991	$898,991

Targeting tumor stress phosphatases for cancer therapy	☑NIH ☐ Other federal (specify:_____ _____) ☐ Nonfederal source (specify: _____)	5/10	$329,946	Pending
	☐NIH ☐ Other federal (specify:_____ _____) ☐ Nonfederal source (specify: _____)		$	$

1. Mitochondrial targeting against radiation damage (Center for Medical Countermeasures Against Radiation renewal)

As part of an NIH application to renew funding for the Center for Medical Countermeasures against Radiation (CMCR), the PI proposed to develop and lead a chemoinformatics core (Core H) of the CMCR. The computational methods developed within the scope of this PA-funded project will allow Core H investigators to: (1) contribute computational expertise to the hit-to-lead activities, library design, lead optimization, and identification of potential mitigators of radiation damage at the CMCR; (2) provide computer-aided predictions of drug toxicity and metabolism; and (3) identify target-mitigator interaction pathways and networks to correlate molecular functions to physiological processes that will help in designing safe and efficient mitigators of radiation damage.

2. Targeting tumor stress phosphates for cancer therapy (grant application)

Tumor cell stress plays an important role in the malignant phenotype and in the cellular response to curative therapy. A major cause of stress in developing tumors is an inadequate blood supply, which leads to tumor hypoxia and nutrient deprivation. DNA damaging agents and irradiation are additional determinants of stress that can be genotoxic and lead to anticancer drug resistance. Cancer cells adapt to these stresses by expressing proteins that help maintain the tumorigenic state through a process that has been termed “non-oncogene addiction” (NOA). We proposed to investigate the role of mitogen-activated protein kinase phosphatases (MKPs) as mediators of adaptive tumor cell survival under conditions of cellular stress. MKPs are dual-specific protein phosphatases that dephosphorylate and inactivate mitogen-activated protein kinases (MAPKs), which have a pivotal role in mitogenic signal transduction, survival, stress response, and programmed cell death. The best-studied member of the MKP family is DUSP1/MKP-1/CL100. While not a transforming oncogene, MKP-1 is overexpressed in many human tumors, and its expression correlates with poor prognosis and decreased patient survival. MKP-1 protects cells from various apoptotic stimuli, limits the efficacy of clinically used antineoplastic agents, and its overexpression is an adaptive response to cisplatin therapy in ovarian cancer. Recently, a large-scale genomic approach sponsored by NIH (the Cancer Genome Atlas, Rembrandt database) has identified that MKP-2 and MKP-3 are overexpressed in gliomas, and their expression correlates with decreased survival.

MKPs are understudied and unexplored as drug targets because potent and specific small molecule inhibitors are lacking; therefore, the armamentarium to investigate MKP biology has been limited to genetic approaches. We recently identified two chemotypes with MKP inhibitory activity *in vitro* and *in vivo*. Both molecules have antiproliferative and antimigratory properties and their biological activity cannot be overridden by MKP-inducing, cytoprotective stimuli. Collectively, these studies form the basis of our overall hypothesis that MKPs, which under physiological conditions attenuate mitogenic and stress signaling, are NOA targets that allow tumors to survive under conditions of cellular stress. We posited that small-molecule inhibition of MKPs may represent a novel pharmacologic strategy to impair the adaptive response of tumors to cellular stress, thereby decreasing cancer cell survival. Accordingly, our specific aims were:

Specific Aim #1. Define potential small molecule binding sites on MKP-1, MKP-2, and MKP-3 and identify the chemical characteristics required for inhibition

Specific Aim #2. Delineate mechanistic and functional consequences of MKP induction or perturbation under conditions of tumor stress

Specific Aim #3. Test the hypothesis that MKP activity mediates cell survival, migration, and invasiveness in cellular and zebrafish models of human glioma

These studies will answer several important questions: (1) What are the determinants that allow MKPs help maintain the tumorigenic state? (2) What aspects of tumor cell maintenance do they control—survival, cell motility, or angiogenesis? (3) Is their cytoprotective activity limited to those MKPs that preferentially regulate death signaling? (4) Will it be possible to exploit MKPs as rational drug targets? The findings will exert a powerful, sustained influence on the field of targeted antineoplastic therapies.

11(B) Are you planning to apply for additional funding in the future to continue or expand the research?

Yes__X_____ No_____

If yes, please describe your plans:

Additional NIH funding will be sought to continue this research.

12. Future of Research Project. What are the future plans for this research project?

We will conduct computational screening of drug-like molecule libraries to identify MKP-1 and -3 inhibitors. We have generated one-microsecond long molecular dynamics (MD) trajectories for MKP-3 and other closely-related MKPs. Careful investigation of these trajectories will enable us to comprehensively characterize the dynamics of the MKP phosphatase domain. Results will be published. We have also investigated and improved a recently proposed simulation-based druggability assessment technique whereby a target protein is simulated in a mixture of water and organic solvents (e.g., isopropanol, acetamide) that mimic the behavior of drug-like molecules. The method enables the identification of new

potential inhibition sites, in addition to assessing the maximal affinity that a drug-like molecule can achieve at a given site.

13. **New Investigator Training and Development**. Did students participate in project supported internships or graduate or post-graduate training for at least one semester or one summer?

Yes____X_____ No_____

If yes, how many students? Please specify in the tables below:

	Undergraduate	Masters	Pre-doc	Post-doc
Male			1	
Female			1	
Unknown				
Total			**2**	

	Undergraduate	Masters	Pre-doc	Post-doc
Hispanic				
Non-Hispanic			2	
Unknown				
Total			**2**	

	Undergraduate	Masters	Pre-doc	Post-doc
White			1	
Black				
Asian			1	
Other				
Unknown				
Total			**2**	

14. **Recruitment of Out-of–State Researchers**. Did you bring researchers into Pennsylvania to carry out this research project?

Yes___X_____ No_____

If yes, please list the name and degree of each researcher and his/her previous affiliation:

Ahmet Bakan, PhD, Koç University, Istanbul, Turkey
Aninditta Dutta, MS, Birla Institute of Technology and Science, India

15. Impact on Research Capacity and Quality. Did the health research project enhance the quality and/or capacity of research at your institution?

Yes____X____ No_____

If yes, describe how improvements in infrastructure, the addition of new investigators, and other resources have led to more and better research:

Software for efficient design of protein-inhibitor complexes has been developed. A new server is also being developed by Dr. Bakan, which will be offered for public use to the scientific community.

16. Collaboration, business and community involvement.

16(A) Did the health research funds lead to collaboration with research partners outside of your institution (e.g., entire university, entire hospital system)?

Yes_____X____ No_____

If yes, please describe the collaborations:

We collaborated with GlaxoSmithKline on this project. Results from this collaboration will be published, and tools will be made publicly available.

16(B) Did the research project result in commercial development of any research products?

Yes_____ No____X_____

If yes, please describe commercial development activities that resulted from the research project:

16(C) Did the research lead to new involvement with the community?

Yes_____ No_____X____

If yes, please describe involvement with community groups that resulted from the research project:

17. Progress in Achieving Research Goals, Objectives and Aims.
List the project goals, objectives and specific aims (as contained in the grant application's strategic plan). Summarize the progress made in achieving these goals, objectives and aims for the entire grant award period. Indicate whether or not each goal/objective/aim was

achieved; if something was not achieved, note the reasons why. Describe the methods used. If changes were made to the research goals/objectives/aims, methods, design or timeline since the original grant application was submitted, please describe the changes. Provide detailed results of the project. Include evidence of the data that was generated and analyzed, and provide tables, graphs, and figures of the data. List published abstracts, poster presentations and scientific meeting presentations at the end of the summary of progress; peer-reviewed publications should be listed under item 20.

This response should be a DETAILED report of the methods and findings. It is not sufficient to state that the work was completed. Insufficient information may result in an unfavorable performance review, which may jeopardize future funding. If research findings are pending publication you must still include enough detail for the expert peer reviewers to evaluate the progress during the course of the project.

Health research grants funded under the Tobacco Settlement Act will be evaluated via a performance review by an expert panel of researchers and clinicians who will assess project work using this Final Progress Report, all project Annual Reports and the project's strategic plan. After the final performance review of each project is complete, approximately 12-16 months after the end of the grant, this Final Progress Report, as well as the Final Performance Review Report containing the comments of the expert review panel, and the grantee's written response to the Final Performance Review Report, will be posted on the CURE Web site.

There is no limit to the length of your response. Responses must be single-spaced below, no smaller than 12-point type. If you cut and paste text from a publication, be sure symbols print properly, e.g., the Greek symbol for alpha (α) and beta (ß) should not print as boxes (□) and include the appropriate citation(s). DO NOT DELETE THESE INSTRUCTIONS.

Specific Aim 1. To identify recurrent features and specific residues playing a key role in the inhibition of selected dual-specificity phosphatases (DUSPs), toward designing potent and selective inhibitors

We previously published a paper describing the discovery of an allosteric inhibitor of DUSP6 (MKP-3). We have further investigated the dynamics and druggability of DUSP6 using simulation methods and have identified a new potential site. Our results suggest that it is possible to identify allosteric inhibitors with better potency and selectivity properties, and we discuss these results in detail below. An RO1 grant application has been submitted to NIH to fund the proposed screening and testing research based on these results.

Specific Aim 2. To assess the entropic contribution to binding free energy by sampling a representative ensemble of protein-ligand complex conformations using elastic network models

We previously published a manuscript showing the importance of protein entropic factors in binding of drug/inhibitor molecules (Bakan and Bahar, 2009). We further investigated and improved a simulation-based method for binding site identification and druggability assessment,

which is discussed at the end of this section. This method had direct impact on our DUSP inhibitor discovery research.

Investigation of DUSP6 dynamics and druggability

An allosteric inhibitor of DUSP6. All tyrosine phosphatases share a common mechanism of dephosphorylation involving highly conserved residues: a catalytic cysteine temporarily accepts the phosphate group from the substrate, an arginine stabilizes the transition state, and a proton donated by an acidic residue in the highly flexible "general acid loop" (GAL) facilitates the hydrolysis of the intermediate phosphothioester. This sequence is called the "catalytic triad," which in DUSP6 is formed by C293, R299, and D262 (Figure 1A). In substrate-inducible DUSPs, such as DUSP6, D262 must be sufficiently close to the catalytic Cys for full activation to occur (Rigas et al., 2001). This activation is presumably achieved by a suitable change in the GAL conformation. Using zebrafish screens, we previously identified a novel small molecule (BCI) that has no redox-active features and inhibits DUSP6 by a novel mechanism (Molina et al., 2009). Since the crystal structure for DUSP6 has been solved (PDB ID: 1mkp), we initially performed docking simulations to determine potential binding sites for BCI on MKP-3 (Molina et al., 2009). These simulations indicated that BCI would preferentially bind the inactive form of DUSP6 at a site that would prevent the movement of the GAL to assume the functional closed state upon extracellular signal-regulated kinase (ERK) binding (Figure 1A). The model also predicted that BCI might inhibit the inducible phosphatase activity of MKP-3 but would have no effect on its basal activity, which was confirmed experimentally (Molina et al., 2009). The discovery that the activity of DUSPs can be regulated through allosteric mechanisms has opened a new avenue to inhibitor discovery.

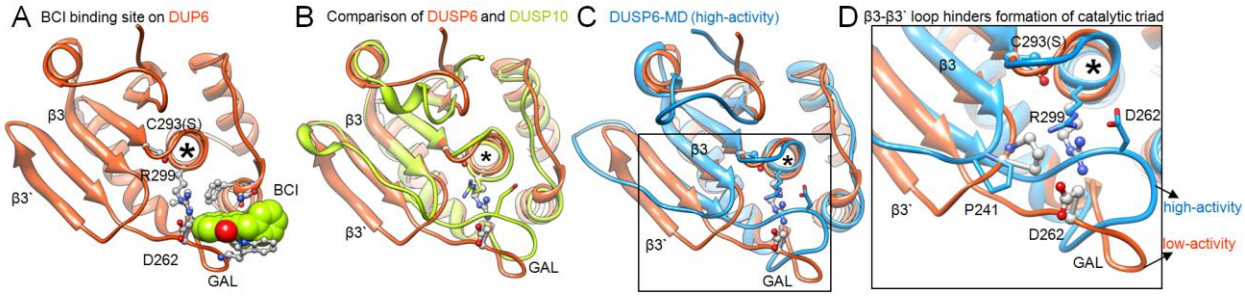

Figure 1. Molecular dynamics simulations of the BCI allosteric site. A. BCI (green) prevents closure of the general acid loop. Upon BCI binding, the GAL containing D262 is held in an open position, locking the enzyme in its low-activity form. The catalytic cavity is marked with an asterisk. B. MKP-3/DUSP6 in its low activity form (orange) compared to constitutively active MKP-5/DUSP10 (green). In MKP-5/DUSP10 the GAL is closed and appears to be coupled to deformation of β3`. C. Comparison of molecular dynamics-derived DUSP6 in its high-activity form (DUSP6-MD, blue) shows β3`deformation similar to MKP-5/DUSP10. D. β3-β3` hairpin loop hinders formation of the catalytic triad. A rigid, structured beta-sheet containing β3-β3` strands joined by a short hairpin keep the GAL in an open position through a proline residue (P241). In the high-activity conformation of DUSP6 (and in constitutively active DUSP10), β3` deforms, P241 moves away, and R299 and D262 reach their active state conformations. This structural transition is presumed to be favored upon ERK binding. Data are from a 30 ns simulation; a snapshot from the 20th ns is shown.

We further investigated the dynamics and druggability properties of the allosteric binding site by performing molecular dynamics (MD) simulations. Because there is no X-ray structure of DUSP6 in its activated state, we initially performed a comparative structural analysis of DUSP6 and DUSP10 as prototypical substrate-activatable and non-activatable DUSPs, respectively. In the low-activity form of DUSP6, (β3`) is part of a rigid beta-sheet that forces the GAL into an open position (Figure 1B, orange). In constitutively active DUSP10, β3` is deformed and the GAL is closed, positioning D262 close to the catalytic cavity (Figure 1B, green). Similar structural changes in the GAL are presumed to happen to DUSP6 upon ERK binding (Fjeld et al., 2000). Our MD simulations with a total time of one microsecond supported this model. In the high-activity form of DUSP6, the GAL was closed and β3` deformed (Figure 1C, blue), similar to DUSP10 (Figure 1B, green). Mechanistically, the model predicts that a rigid, structured beta-sheet containing β3-β3` strands joined by a short hairpin keeps the GAL in an open position through a proline residue (P241). In the high-activity conformation of DUSP6 (and in constitutively active DUSP10), β3` deforms, P241 moves away, and R299 and D262 reach their active state conformations.

A novel small molecule binding site on DUSP6. Based on the above model and our previous experimental results (Molina et al., 2009), we hypothesized that alternative sites that stabilize the conformation of the β3` strand could elicit the allosteric effect triggered by BCI binding. To identify such sites, we analyzed the druggability of the DUSP6 catalytic domain using a first principles-based method (Seco et al., 2009). In this method, the enrichment of isopropanol on particular sites on the protein surface is translated into a druggability index. We calculated the maximal achievable affinity for the BCI binding site to be 3 μM, which is consistent with the BCI's antiphosphatase activity (Molina et al., 2009). We also detected a second site with predicted nanomolar binding affinity located on the strand β5 (Figure 2A). Notably, the reconstruction of the unit cell in the DUSP6 crystal structure (Stewart et al., 1999) shows that DUSP6 makes extensive crystal contacts at this particular region (Figure 2B), in support of the propensity of this site to make intermolecular interactions. Four key residues on the neighboring DUSP6 molecule that make crystal contacts are shown in Figure 2C. Calculation of the buried surface area at this region yielded 139, 122, 97, and 75 $Å^2$ for W264, P334, M337, and L341, respectively. The respective interaction energies were calculated to be -3.0, -4.7, -3.1, and -2.5 kcal/mol, which cumulatively lead to a 0.2 nM binding affinity. This value, comparable to our druggability predictions, provides independent support for targeting this particular pocket.

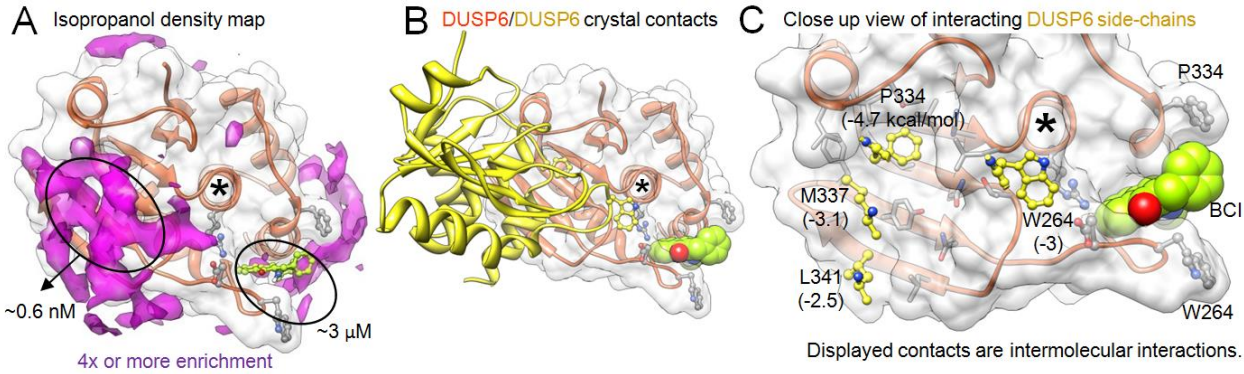

Figure 2. A novel allosteric small molecule binding site on DUSP6. A. Results from druggability simulations. Purple clouds correspond to regions enriched in isopropanol concentration, corresponding to energetically favorable sites for drug-like molecule binding. The BCI site was calculated as a μM site. A more distant site with nM druggability index was also observed. *B. DUSP6 makes extensive crystal contacts at the distant site.* Analysis of DUSP6/DUSP6 interactions (by reconstruction of the unit cell using crystal parameters) showed favorable interactions between monomers (second DUSP6 shown in yellow), suggesting the novel site can accommodate small molecules. *C. Key interacting side-chains on neighboring DUSP6 monomers (shown in yellow) occupy binding pockets at the distant site and constitute anchors for biased docking and virtual screening studies for identification of new chemotypes.* Interaction energies predicted by ANCHOR server are shown in parentheses.

In summary,

- Molecular modeling techniques have predicted a novel allosteric mechanism of inhibition for DUSP6, which was supported by experimental evidence.
- The catalytic domains of DUSP6 have at least two allosteric binding sites for small molecules that could be exploited for small molecule inhibitor discovery.

Explore a new, high-affinity binding pocket on DUSP6 as a novel site for allosteric inhibition. Our druggability calculations and the crystal contacts observed in DUSP6 suggest that the β3` site offers an alternative binding site that may provide a novel point of attack for allosteric inhibition. The size of this nanomolar pocket expands up to 317 $Å^3$, calculated from snapshots taken from druggability simulations using Q-SiteFinder (Laurie and Jackson, 2005). This size is able to accommodate lead-like compounds. The pocket is lined by mostly hydrophobic residues (L225, F253, L292 labeled in Figure 3A, and L228, I236, L290 deep in the pocket; and also Y255, H292, and N238), which is a desirable target site feature for designing drug-like molecules (Hopkins and Groom, 2005).

We plan to further test the hypothesis that the nanomolar site will provide inhibitors of DUSP activity by performing computational docking screens, followed by purchasing hits and their biological evaluation. To identify new leads, we will use contacting side chains as anchors (Rajamani et al., 2004). Our initial screening setup, which is currently being run on 10 processors simultaneously, involves docking with a scaffold match constraint using GOLD software (Jones et al., 1997). The current scaffold is based on the experimental coordinates of a Phe334 side-chain of the neighboring DUSP6 monomer. As shown in Figure 3A, the scaffold

aims to mimic interactions found in the DUSP6 crystal structure. Compounds may extend from any combination of atoms marked by arrows in Figure 3A. To validate our experimental setup, we began to screen a limited subset (50,000) of commercially available lead-like compounds (mw < 350) stored in the ZINC database, which has approximately 3 x 10^6 compounds (Irwin and Shoichet, 2005). Examples of potential ligands are shown in Figure 3B and C. Use of the scaffold match constraint and further filtering with the PB/SA (Poisson-Boltzmann/surface area) scoring method (Grant et al., 2001), which is a more rigorous method that results in higher enrichment rates (Brown and Muchmore, 2006), is expected to provide < 5,000 hits based on published estimation methods (Shoichet et al., 2002). Chemical structures from the virtual screen with favorable binding free energy will be further prioritized by cluster analysis using Leadscope Enterprise Predictive software version 2.4.14-5 (Columbus, OH) as described previously (Johnston et al., 2007). Our collaborators at the University of Pittsburgh Drug Discovery Institute (UPDDI) will use ChemNavigator to identify commercial sources of selected hits and to purchase sufficient amounts to test for cellular and *in vitro* antiphosphatase activity.

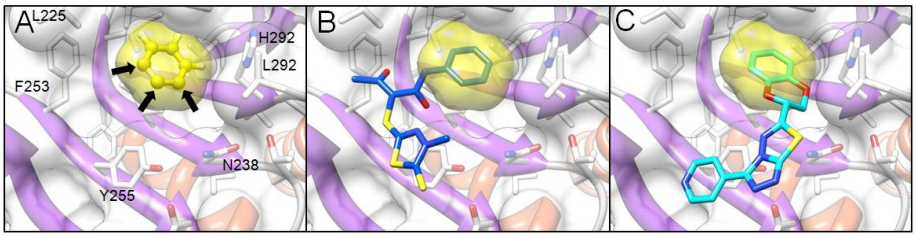

Figure 3. A. Scaffold match constraint based on Phe334 side chain. Compounds are docked if they have a ring structure with matching hydrogen atoms. The scaffold has three extension points at positions marked with an arrow. *B and C. Compounds expand the scaffold from different atoms (arrows in A.)* Compounds from a limited pilot screen of the ZINC database (http://zinc.docking.org) to show examples matched to the scaffold

Druggability evaluation: improved algorithm and expanded coverage

Identification of potential binding sites on a target protein and assessment of their druggability is essential in drug discovery. Recently, Seco and Barril introduced an innovative, first principles-based technique to this aim (Seco et al., 2009). The method, based on the examination of the isopropanol accessibility of the target protein surface in molecular dynamics simulations, is in principle applicable to different types of binding sites, including protein-protein interfaces, catalytic sites, or allosteric sites. Yet, the use of a hydrophobic/polar probe (isopropanol) precludes its applicability to binding sites containing charged residues, such as the active site of protein tyrosine phosphatase 1B. In addition, the intermediate results were not easily interpretable, and the methodology contained redundant assumptions. We proposed: (1) a new methodology based on the same principles, but with simplified assumptions to facilitate the interpretation of the results; (2) a new, fully-automated algorithm that does not require user intervention; and (3) a new probe set (containing isopropanol, isobutane, acetamide, acetate, and isopropylamine) that expands the coverage of the methodology to binding sites with diverse physicochemical properties. We benchmarked the new methodology against a diverse set of binding sites from five different proteins. We also investigated the effect of adopting different probe partial charge sets and probe compositions on the predictive ability of the method.

We also developed tools for fully automated system preparation and druggability analysis. These tools will be released for public use by the end of 2010 with an accompanying journal article. These tools are expected to have an impact on guiding research in targeting proteins with feature poor catalytic sites by helping identification of novel sites. This work was performed in collaboration with GlaxoSmithKline.

Examination of metal-binding sites in proteins as target sites for mediating function
Metal-binding proteins are associated with a variety of cellular functions like transport, cell signaling, cofactors for enzyme catalysis, and many others. With increasing structural data on this group, we have systematically analyzed the collective dynamics of a representative set of metal-binding proteins. We have tried to ascertain whether metal-binding sites have distinguishing features or positions that enable them to achieve their role in the activity of the proteins that bind them. We: (1) use elastic network models (Gaussian Network Model) to see how metal-binding sites contribute to global dynamics, (2) analyze their solvent accessibility, and (3) use Markovian stochastic analysis to elucidate their signal transduction properties and their involvement in allosteric communication.

Analyzing the global mobility profiles of 30 *apo-holo* pairs illustrated the insensitivity of global modes to the presence/absence of bound metal. Hence we chose 145 *holo* structures and examined whether metal-binding residues (cysteines (C), histidines (H), glutamates (E) and aspartates (D) or "CHED") occupy key positions (near hinges) and, thus, appear as minima in the global mode profile (Figure1). The histogram in Panel A shows that metal-binding CHED residues, indicated by red bars, have a clear bias toward lower mobilities compared to all residues, as well as their non-metal-binding counterparts. A closer analysis of the mobilities of each of the subsets showed that metal-coordinating histidine, glutamate, and aspartate (HED) prefer to be present in regions of lower mobility, while cysteine shows the opposite trend.

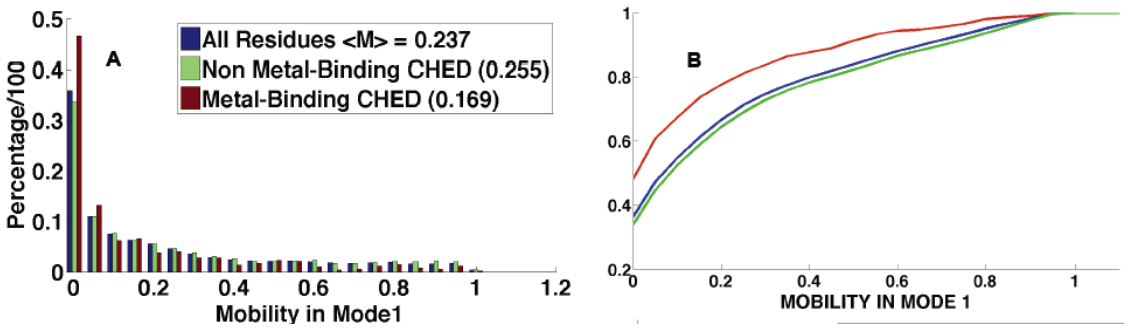

Figure 1. Shows the mobility in Global Mode 1. A: displays the histograms of mobilities for the three subsets. For comparative purposes, the mobilities were normalized in the range [0, 1] for each protein. B: Shows the cumulative distribution for three subsets

A comparative analysis of solvent accessibility of the two subgroups of CHED residues shed some light on their distinctive properties. We oberved a predisposition of metal-coordinating HED to occupy regions with limited, if any, exposure to solvent, in spite of being polar or charged. Thus, low solvent accessibility and low mobility can be used as "features" in facilitating

the identification of potentially metal-coordinating HED residues in putative metal-binding proteins.

We previously developed a Markov model for describing the stochastics of signal transmission in proteins modeled as networks of nodes and springs (Chennubhotla and Bahar, 2007). The two main metrics of communication propensities are: (1) hitting time, *H(i,j)*, which represents an average number of steps for signals to be transmitted from residue *j* to residue *i* and (2) commute time, *C(i,j) = H(i,j) + H(j,i)*. These properties for a representative metal-binding protein have been shown in Figure 2 (A & B). Panel C shows the average hit time versus residue number, evaluated from the average of *H(i,j)* over all starting points *j*.
Red markers highlighting metal-binding residues show that they occur at minima, pointing to efficient signal transmission. Panel D shows the average hit-time versus variance for each residue. The metal-binding residues identified by red markers show that they have fast, as well as precise, communication, as shown by small standard deviation.

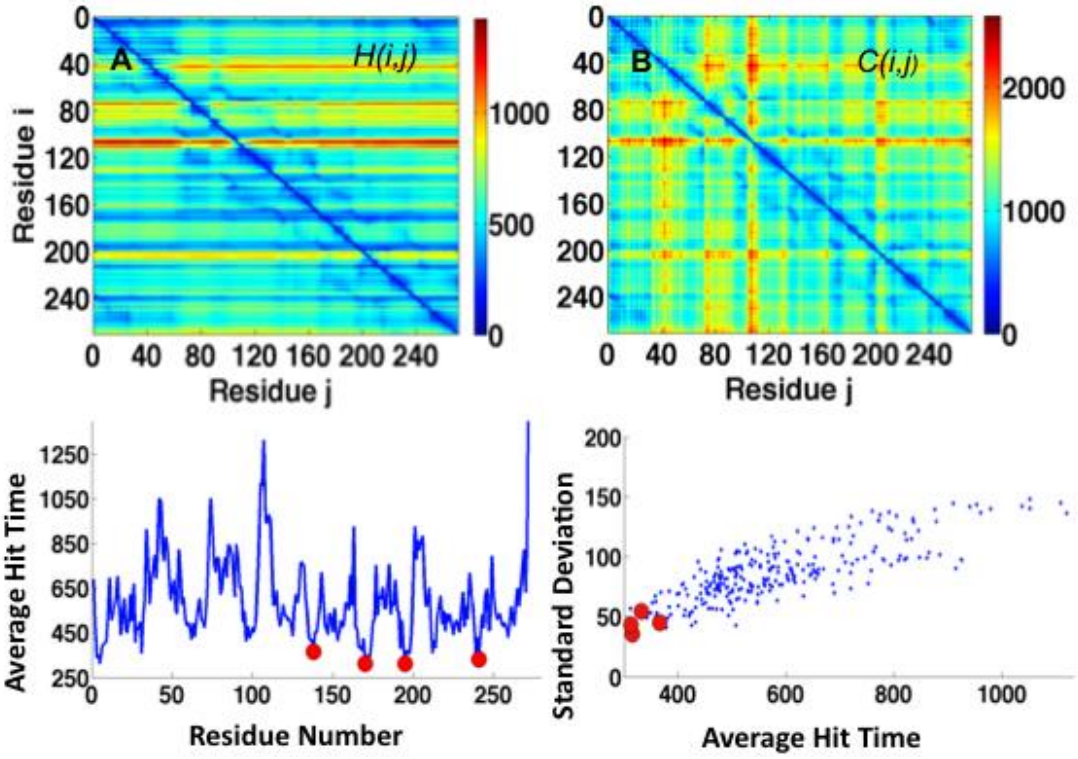

Figure 2. Signal propagation properties illustrated for a Zn2+ binding protein (1I6N). A: Hitting time B: Commute time C: Residue number versus average hitting time with red markers showing metal-binding sites that show minimal hit times. D: Average hit times versus standard deviation. Metal binding sites are made prominent with red markers to highlight minimal variance.

A comprehensive analysis of the average hit time versus standard deviation of residues participating in metal-binding in all 145 *holo* structures reiterated the above trend, where almost all metal-binding residues showed minimal hit times and minimal variance, pointing toward their fast and precise signal transmission properties.

Thus, our results show that metal-binding sites may have an inherent preference to undergo minimal fluctuations in their positions, which might enable them to elicit cooperative responses, occupy central/buried positions (despite being polar or charged), and possess unique signal transduction properties that suggest that these sites may be selected to achieve optimum allosteric communication. These properties give us useful insights into the basic principles for designing metal-binding sites.

Poster Presentations and Talks

Aforementioned results and methodological developments were presented at the University of Pittsburgh's Department of Computational and Systems Biology annual retreat, and the presentation was awarded the prize for best presentation. Results were also demonstrated at the University of Pittsburgh Postdoctoral Association annual meeting as a poster by Dr. Bakan; the poster was chosen as one of the best 10 posters at the meeting. A presentation will be given at the Zing conference on Structural Biology and Drug Discovery in Cancun, Mexico, in December, 2010.

References

Bakan A, Bahar I. The intrinsic dynamics of enzymes plays a dominant role in determining the structural changes induced upon inhibitor binding. *Proc Natl Acad Sci USA* 2009, 106, 14349.

Brown SP, Muchmore SW. High-throughput calculation of protein-ligand binding affinities: modification and adaptation of the MM-PBSA protocol to enterprise grid computing. *J Chem Inf Model*. 2006;46(3):999-1005.

Fjeld CC, Rice AE, Kim Y, Gee KR, Denu JM. Mechanistic basis for catalytic activation of mitogen-activated protein kinase phosphatase 3 by extracellular signal-regulated kinase. *J Biol Chem*. 2000;275(10):6749-57.

Grant JA, Pickup BT, Nicholls A. A smooth permittivity function for Poisson-Boltzmann solvation methods. *J Comput Chem*. 2001;22(6):608-40.

Hopkins AL, Groom CR. Target Analysis: a priori Assessment and Drugability. In: Waldmann H, Koppitz M, editors. Small molecule--protein interactions: Springer Science + Business Media, Heidelberg; 2003.

Irwin JJ, Shoichet BK. ZINC--a free database of commercially available compounds for virtual screening. *J Chem Inf Model*. 2005;45(1):177-82. PMCID: PMC1360656.

Johnston PA, Foster CA, Shun TY, Skoko JJ, Shinde S, Wipf P, et al. Development and implementation of a 384-well homogeneous fluorescence intensity high-throughput screening

assay to identify mitogen-activated protein kinase phosphatase-1 dual-specificity protein phosphatase inhibitors. *Assay Drug Dev Technol*. 2007;5(3):319-32.

Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol*. 1997;267(3):727-48.

Laurie AT, Jackson RM. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics*. 2005;21(9):1908-16.

Lazo JS, Skoko JJ, Werner S, Mitasev B, Bakan A, Koizumi F, et al. Structurally unique inhibitors of human mitogen-activated protein kinase phosphatase-1 identified in a pyrrole carboxamide library. *J Pharmacol Exp Ther*. 2007;322(3):940-7.

Molina G, Vogt A, Bakan A, Dai W, de Oliveira PQ, Znosko W, et al. Zebrafish chemical screening reveals an inhibitor of Dusp6 that expands cardiac cell lineages. *Nat Chem Biol*. 2009;6(8):680-7. PMCID: PMC2771339.

Rajamani D, Thiel S, Vajda S, Camacho CJ. Anchor residues in protein-protein interactions. *Proc Natl Acad Sci U S A*. 2004;101(31):11287-92. PMCID: PMC509196.

Rigas JD, Hoff RH, Rice AE, Hengge AC, Denu JM. Transition state analysis and requirement of Asp-262 general acid/base catalyst for full activation of dual-specificity phosphatase MKP3 by extracellular regulated kinase. *Biochemistry*. 2001;40(14):4398-406.

Seco J, Luque FJ, Barril X. Binding site detection and druggability index from first principles. *J Med Chem*. 2009;52(8):2363-71.

Shoichet BK, McGovern SL, Wei B, Irwin JJ. Lead discovery using molecular docking. *Curr Opin Chem Biol*. 2002;6(4):439-46.

Stewart AE, Dowd S, Keyse SM, McDonald NQ. Crystal structure of the MAPK phosphatase Pyst1 catalytic domain and implications for regulated activation. *Nat Struct Biol*. 1999;6(2):174-81.

Chennubhotla C, Bahar I. Signal propagation in proteins and relation to equilibrium fluctuations. *PLoS Comput Biol*. 2007;1716-1726.

Dutta A, Bahar, I. (2010) Metal-binding sites are designed to achieve optimal mechanical and signaling properties. *Structure* (Accepted).

18. Extent of Clinical Activities Initiated and Completed. Items 18(A) and 18(B) should be completed for all research projects. If the project was restricted to secondary analysis of clinical data or data analysis of clinical research, then responses to 18(A) and 18(B) should be "No."

18(A) Did you initiate a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
___X__No

18(B) Did you complete a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
___X__No

If "Yes" to either 18(A) or 18(B), items 18(C) – (F) must also be completed. (Do NOT complete 18(C-F) if 18(A) and 18(B) are both "No.")

18(C) How many hospital and health care professionals were involved in the research project?

_____Number of hospital and health care professionals involved in the research project

18(D) How many subjects were included in the study compared to targeted goals?

_____Number of subjects originally targeted to be included in the study
_____Number of subjects enrolled in the study

18(E) How many subjects were enrolled in the study by gender, ethnicity and race?

Gender:
_____Males
_____Females
_____Unknown

Ethnicity:
_____Latinos or Hispanics
_____Not Latinos or Hispanics
_____Unknown

Race:
_____American Indian or Alaska Native
_____Asian
_____Blacks or African American
_____Native Hawaiian or Other Pacific Islander
_____White
_____Other, specify:_____
_____Unknown

18(F) Where was the research study conducted? (List the county where the research study was conducted. If the treatment, prevention and diagnostic tests were offered in more than one county, list all of the counties where the research study was conducted.)

19. Human Embryonic Stem Cell Research. Item 19(A) should be completed for all research projects. If the research project involved human embryonic stem cells, items 19(B) and 19(C) must also be completed.

19(A) Did this project involve, in any capacity, human embryonic stem cells?
_____Yes
___X_ No

19(B) Were these stem cell lines NIH-approved lines that were derived outside of Pennsylvania?
_____Yes
_____ No

19(C) Please describe how this project involved human embryonic stem cells:

20. Articles Submitted to Peer-Reviewed Publications.

20(A) Identify all publications that resulted from the research performed during the funding period and that have been submitted to peer-reviewed publications. Do not list journal abstracts or presentations at professional meetings; abstract and meeting presentations should be listed at the end of item 17. **Include only those publications that acknowledge the Pennsylvania Department of Health as a funding source** (as required in the grant agreement). List the title of the journal article, the authors, the name of the peer-reviewed publication, the month and year when it was submitted, and the status of publication (submitted for publication, accepted for publication or published.). Submit an electronic copy of each publication, listed in the table, in a PDF version 5.0.5 format, 1,200 dpi. Filenames for each publication should include the number of the research project, the last name of the PI, the number of the publication and an abbreviated research project title. For example, if you submit two publications for PI Smith for the "Cognition and MRI in Older Adults" research project (Project 1), and two publications for PI Zhang for the "Lung Cancer" research project (Project 3), the filenames should be:

Project 1 – Smith – Publication 1 – Cognition and MRI
Project 1 – Smith – Publication 2 – Cognition and MRI
Project 3 – Zhang – Publication 1 – Lung Cancer
Project 3 – Zhang – Publication 2 – Lung Cancer

If the publication is not available electronically, provide 5 paper copies of the publication.

Note: The grant agreement requires that recipients acknowledge the Pennsylvania Department of Health funding in all publications. Please ensure that all publications listed

acknowledge the Department of Health funding. If a publication does not acknowledge the funding from the Commonwealth, do not list the publication.

Title of Journal Article:	Authors:	Name of Peer-reviewed Publication:	Month and Year Submitted:	Publication Status (check appropriate box below):
Metal-binding sites are designed to achieve optimal mechanical and signaling properties	Dutta A., and Bahar, I	Structure	Feb, 2010	☐Submitted ☑Accepted ☐Published
				☐Submitted ☐Accepted ☐Published
				☐Submitted ☐Accepted ☐Published

20(B) Based on this project, are you planning to submit articles to peer-reviewed publications in the future?

Yes____X___ No_____

If yes, please describe your plans:

We have generated more than one microsecond of dynamics data on DUSP6 and other relevant DUSPs. These data enable understanding the intrinsic dynamics and activation of DUSPs and are useful in modeling interactions with inhibitors. We plan to publish these results in a journal article. We have also made significant improvements in a simulation-based druggability assessment method and developed tools for automated system setup and analysis. A manuscript describing the improvements in the theory will be submitted in fall 2010. Another manuscript describing the tools will also be submitted at the end of 2010. Tools will be released publicly at the same time.

Results from computational screening of inhibitors for the newly discovered allosteric site will also be presented in peer-reviewed journals after we have obtained experimental results from our collaborators at UPDDI.

21. Changes in Outcome, Impact and Effectiveness Attributable to the Research Project. Describe the outcome, impact, and effectiveness of the research project by summarizing its impact on the incidence of disease, death from disease, stage of disease at time of diagnosis, or other relevant measures of outcome, impact or effectiveness of the research project. If there were no changes, insert “None”; do not use “Not applicable.” Responses must be

single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

We have identified a novel allosteric inhibitor of DUSP6, which has opened new venues for inhibition of DUSPs. Further investigation of DUSP dynamics and druggability resulted in the discovery of a new site, offering ability to control DUSP activity allosterically by potent inhibitors. Our research also showed that flexibility of drug target proteins can be modeled using efficient analytical methods and resulted in tools for the use of the scientific community community free of charge.

22. Major Discoveries, New Drugs, and New Approaches for Prevention Diagnosis and Treatment. Describe major discoveries, new drugs, and new approaches for prevention, diagnosis and treatment that are attributable to the completed research project. If there were no major discoveries, drugs or approaches, insert "None"; do not use "Not applicable." Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

The discovery of a new allosteric inhibitor and sites for DUSPs is a major achievement in the field of drug discovery. These findings will enable development of potent and specific inhibitors for interrogation of the role of DUSPs in development and progression of human diseases.

23. Inventions, Patents and Commercial Development Opportunities.

23(A) Were any inventions, which may be patentable or otherwise protectable under Title 35 of the United States Code, conceived or first actually reduced to practice in the performance of work under this health research grant? Yes_____ No___X_____

If "Yes" to 23(A), complete items a – g below for each invention. (Do NOT complete items a - g if 23(A) is "No.")

a. Title of Invention:

b. Name of Inventor(s):

c. Technical Description of Invention (describe nature, purpose, operation and physical, chemical, biological or electrical characteristics of the invention):

d. Was a patent filed for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes_____ No____

If yes, indicate date patent was filed:

e. Was a patent issued for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes_____ No____

If yes, indicate number of patent, title and date issued:
Patent number:
Title of patent:
Date issued:

f. Were any licenses granted for the patent obtained as a result of work performed under this health research grant? Yes_____ No___

If yes, how many licenses were granted?_____

g. Were any commercial development activities taken to develop the invention into a commercial product or service for manufacture or sale? Yes__ No___

If yes, describe the commercial development activities:

23(B) Based on the results of this project, are you planning to file for any licenses or patents, or undertake any commercial development opportunities in the future?

Yes_____ No____X _____

If yes, please describe your plans:

24. Key Investigator Qualifications. Briefly describe the education, research interests and experience and professional commitments of the Principal Investigator and all other key investigators. In place of narrative you may insert the NIH biosketch form here; however, please limit each biosketch to 1-2 pages. *For Nonformula grants only – include information for only those key investigators whose biosketches were not included in the original grant application.*

IVET BAHAR, PhD

(a) Professional Preparation
Bogazici University, Istanbul, Turkey, Chemical Engineering, BS, 1980
Bogazici University, Istanbul, Turkey, Chemical Engineering, MS, 1983
Istanbul Technical University, Turkey, Chemistry, PhD, 1986

(b) Appointments

Since 2010	Professor and John K. Vries Chair Department of Computational & Systems Biology, School of Medicine, University of Pittsburgh
Since 2006	Faculty, Drug Discovery Institute School of Medicine, University of Pittsburgh
2005-2009	Founding Director at U Pitt, PhD Program in Computational Biology, Carnegie Mellon University (CMU) and University of Pittsburgh (CMU Co-Founder/Director: R.F. Murphy)

2004-2010	Founding Chair, Computational Biology Department School of Medicine, University of Pittsburgh
2001-2004	Founding Director, Center for Computational Biology & Bioinformatics School of Medicine, University of Pittsburgh W1041 Biomedical Science Tower (BST)
2001-2004	Professor, Molecular Genetics & Biochemistry School of Medicine, University of Pittsburgh E1240 BST
1998-2001	Director, Turkish National Scientific and Technological Research Council (TUBITAK) Materials Research Unit, Istanbul, Turkey
1994-2000	Visiting Scientist, Laboratory of Experimental & Computational Biology, Division of Basic Sciences, National Cancer Institute, NIH, Bethesda, MD (summer visits)
1993-2001	Professor, Chemical Engineering Department, Bogazici University, Istanbul
1992-2000	Co-Founding Director, Polymer Research Center, Bogazici University, Istanbul
1992-1993	Fogarty Fellow, Laboratory of Experimental & Computational Biology, Division of Basic Sciemce, NIH (summer)
1989-1998	Visiting Scientist, Ecole Supérieure de Physique et Chimie, Laboratoire de Physico-chimie et Macromoléculaire, Paris, France (several short-term appointments)
1988-1991	Visiting Associate and UNESCO Fellow, Institute of Polymer Science, U of Akron, OH
1987-1993	Associate Professor, Chemical Engineering Department, Bogazici University
1986-1987	Assistant Professor, Chemical Engineering Department, Bogazici University
1984-1986	Research Assistant, Chemistry, School of Arts & Sciences, Istanbul Technical University

(c) Publications

- Yang LW, Eyal E, Bahar I, Kitao A. (2009) Principal component analysis of native ensembles of biomolecular structures (PCA_NEST): insights into functional dynamics. *Bioinformatics* 25: 606-614 PMID: 19147661
- Bakan A, Bahar I. (2009) The intrinsic dynamics of enzymes plays a dominant role in determining the structural changes induced upon inhibitor binding. *Proc Natl Acad Sci USA* **106**, 14349-54 PMID: 19706521
- Shrivastava IH, Jiang J, Amara SG, Bahar I. (2008) Time-resolved mechanism of extracellular gate opening and substrate binding in glutamate transporter. *J. Biol. Chem.* 283, 28680-90, PMID: 18678877.
- Eyal E, Yang LW, Bahar I. (2006) "Anisotropic network model: systematic evaluation and a new web interface. *Bioinformatics* **22**: 2619-27. PMID: 16928735
- Yang LW, Liu X, Jursa CJ, Holliman M, Rader AJ, Karimi H, Bahar I. (2005) "iGNM: A database of protein functional motions based on Gaussian Network Model *Bioinformatics* 21: 2978-87 PMID: 15860562
- Bahar I, Lezon TR, Yang LW, Eyal E. (2010) Global dynamics of proteins: bridging between structure and function. *Annu Rev Biophys.* 39: 23-42.
- Bahar I, Lezon T, Bakan A, Shrivastava I. (2010) Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. *Chem. Rev.* 110, 1463-1497. PMID: 19785456

- Yang Z, Májek P, I Bahar. (2009) Allosteric transitions of supramolecular systems explored by network models: application to chaperonin GroEL. *PLoS Comput Biol* 5(4): e1000360. PMID: 19381265
- Bahar I, Chennubhotla C, Tobi D. (2007) Intrinsic dynamics of enzymes in the unbound state and relation to allosteric regulation. *Curr Opin Struct Biol.* **17**: 633-40. PMID: 18024008
- Chennubhotla C, Bahar I: (2006) Markov propagation of allosteric effects in biomolecular systems: application to GroEL-GroES. *Molecular Systems Biology* 2: 36-48 PMID: 16820777

(d) Synergic Activities

1. Founder of University of Pittsburgh School of Medicine Department of Computational Biology (2004) and Center for Computational Biology & Bioinformatics (2001); co-founder of Polymer Research Center at Bogazici University (1992).
2. Founding director of joint PhD program in computational biology, between the University of Pittsburgh and Carnegie Mellon University, which received an HHMI Interfaces award in 2005 and a National Institute of Biomedical Imaging and Bioengineering T32 in 2009. Play a leading role in the curriculum design and preparation and teaching of several new courses in computational biology and bioinformatics for undergraduate and graduate students (1990-2005), including a new course in structural bioinformatics (2002) and a new molecular biophysics III course (2004) for School of Medicine PhD students.
3. Active role in several international scientific societies and activities (as executive committee member of the Biophysical Society, elected member of European Molecular Biology Organization (EMBO), member of Turkish Academy of Sciences, editorial activities, organization of workshops/meetings).
4. Director of the NSF/NIH Bioengineering and Bioinformatics Summer Institute (BBSI) on "Simulation and Computer Visualization of Biological Systems at Multiple Scale" –a joint effort between the University of Pittsburgh, Carnegie Mellon University, Duquesne University, and Pittsburgh Supercomputing Center (2002-2009).
5. Development and implementation of iGNM, ANM, and PCA_NEST servers and databases for protein dynamics.