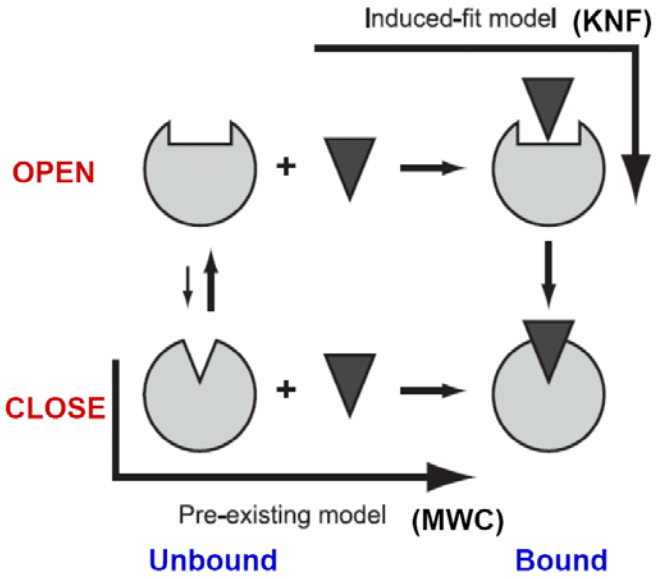
Research Activities and Findings in the past
& Ongoing Research Projects

Biomolecules operate as nanomachines to carry out precise biochemical functions by incurring proper conformational changes. To realize the spatial extent, the time scale and the *physical origin* of such motions is of the utmost importance and also of my great interest. As a theoretical biophycist and a graduate from school of medicine, I exploit theoretical/computational means as well as develop new methodologies to gain insights of life in the molecular level. MD simulations, Normal Mode Analysis using standard forcefields, Elastic Network Model and Gaussian Network Model are tools that I used to study the equilibrium conformational dynamics of proteins and nucleotides in all-atom (fine-grained) or in residue (coarse-grained) levels.

Conformational changes of biomolecules occur to accommodate substrate or to facilitate interactions with their binding partners form the simplest dynamic picture of life in the molecular level. KNF (induce-fit) and MWC (pre-existing equilibrium) models were suggested decades ago to explain the allosteric phenomenon of hemoglobin. However, the physical origin(s) of the two models has not been understood in enough depth at the molecular level so as to properly explain the observed conformational changes.

At physiological temperature and pressure, biomolecules have their intrinsic motions about the mean. A *pre-existing state*, deviating from the mean at elevated energy, could favor a binding event, either by a ligand or other protein partners [24]. On the other hand, an incoming ligand, landing in the open form of the receptor, could induce a closure of protein domains resulting in a binding event. Via which one of the two seemingly different pathways for free proteins to migrate into bound complexes highlights a constant debate for decades.



My past researches had a heavy emphasis on the "pre-existing" subject while during the past two years I took up the "induce-fit" route and found these two models (MWC&KNF) could basically reconcile at a simple yet carefully formulated theory, Linear Response Theory (LRT), recently revisited by theoretical biophysicists [23]. With Drs Nobuhiro Gō and Akio Kitao, we went one step further to formulate a time-dependent LRT to study primary response of myoglobin upon photodissociation.

The **ongoing** projects and **published** research results of mine are listed as follows.

Prediction of catalytic rate of enzyme mutants and mechanical reasons for some evolutionarily conserved residues – theory and experiment (with Dr Shakhnovich) (ongoing)

Allosteric regulation in enzyme activity can be seen as the ability of proteins to visit many conformational states. Mechanical energy is required for protein to take from one state to the other and the binding/hydrolysis of a cognate substrate can provide such energy. We hypothesize that there are some key mechanical residues (KMR), evolutionarily conserved, which could perturb the energy dramatically upon mutation, and therefore possibly jeopardize enzyme activity from places remote to the catalytic center. People have shown experimentally that single mutation G121V of βF-βG loop, which is more than 19Å away from the closest atom of DHF at the active site, reduces the rate of hydride transfer in DHFR by 163 fold. Also, double mutation M42F-G121A impairs enzyme activity by 21 times more than individual effects combined would do (i.e. highly non-additive) when the two residues are 20Å away from each other. We would first apply more comprehensive, published approaches that are to use **classic + quantum mechanics simulations** to predict mutant activities. Then, we will attempt to develop **purely classic-mechanics-based simple models** to tackle the same issue. Shall this be successful, we will then apply the models to other enzymatic systems to verify our first-principal-based understandings of enzyme catalytic mechanism. We also have **a wet lab to verify our predictions** in a uniformed, controlled environment. This theory-joint-experiment project will first identify a few KMR and then theoretically predict by how much the catalytic rate would reduce/increase if the possibility for DHFR to visit a *catalysis-prompted bound state* is impaired/boosted by a given mutation.

Time-dependent Linear Response Theory to studying primary responses of myoglobin (with Drs Nobuhiro Gō and Akio Kitao) (ongoing/submitted [14])

The primary response of proteins occurs at femto- to picosecond range is understood as protein response to force perturbations. The extent of response at a given atom *i* can be expressed as the sum of component *ij* of fluctuation variance multiplying force f_j applied on atom *j* [23].

In the past year, we have formulated a time-dependent response function

$$< \Delta\vec{r}_i(t) >_f = \frac{1}{k_B T} \int_0^t dt' \sum_{j=1}^{m} < \Delta\dot{\vec{r}}_j(0) \Delta\vec{r}_i(t') >_0 \; f_j(t-t')$$

,within which the time-dependent covariance matrix (or so derived matrices) can be expressed in the normal mode space with solvent friction taken into account [22] (friction is obtained from velocity time correlation function, calculated from MD trajectory). Proteins, projected into each normal mode, as a single oscillator subject to solvent friction and random forces, can be traced to its time-dependent conformational changes upon external or internal perturbations (e.g photodissociation to release otherwise bound monoxide ligand). We study the time course of structural changes between deoxy- and carbonmonoxy myoglobin (CO-Mb) upon multi- or single point (at Fe of heme) force perturbations and compare the results with observables of UV resonance raman (UVRR) spectrum upon photodissociation of CO. A qualitative and quantitative agreement has

been obtained for the relaxation and primary-response time constants at respectively 10s picoseconds and 100s femtoseconds to a few picoseconds.

Principal component analysis of native ensembles of biomolecular structures (PCA_NEST): insights into functional dynamics (with Drs Ivet Bahar and Akio Kitao) (published [13])

The PCA of the native ensemble of structures accessible to a given biomolecule is shown in the present study to provide information on functional dynamics and biomolecular design features. Calculations performed for NMR ensembles of structures show that, in 20 out of 24 examined proteins, the conformational differences $\Delta\{\Phi\}$ observed between the NMR structures and their X-ray counterparts are consistent with the principal modes of motion identified by PCA and supported by the physics-based anisotropic network model (ANM). These 'robust modes', insensitive to model parameters and methods, point to directions of conformational changes that are energetically favored by the particular topology of native contacts.

We argue that the molecules are designed to so move for functional reasons. Evidence is given here that catalytic residues of 12 examined enzymes, structures of which are determined by NMR, are shown to occupy highly constrained, 'immobile', positions along the dominant PC profiles. Notably, the residues identified to be highly constrained in these robust PC modes are verified to be evolutionarily conserved, further supporting the utility of the PCA of structural ensembles for identifying *potential functional sites* (PFSs).

In the supplementary material of this article, we reported a study that shows a handful of conformers indicated by low PCA modes have been sufficient to reproduce the NMR order parameters. The result shows that to have a proper sampling of conformational space, rather than the size of the ensemble, is the key to successfully reproduce the experimentally characterized observables.

Finally, we compile the present mathematical tools into public accessible domain. The online portal, **PCA_NEST** (**PCA** of **N**ative **E**nsembles of **ST**ructures), uses as input ensembles of protein and/or polynucleotide structures accessible under native state conditions, and releases the corresponding principal modes of structural changes and potential functional sites (PFSs) via a user-friendly interface. PCA_NEST is at http://ignm.ccbb.pitt.edu/oPCA_Online.htm

Insights into Equilibrium Dynamics of Proteins from Comparison of NMR and X-ray Data with Computational Predictions (with Drs Ivet Bahar and Angela Gronenborn) (published [12])

PCA_NEST work described above is a continuous effort to answer the question whether or not positional uncertainties of NMR structures and the structural differences between Xray-determined conformers for the same proteins are of "errors" in structural determination or that they do carry information of functional dynamics. In 2007, we presented a study that the positional uncertainties of NMR structures do outline the protein equilibrium dynamics that physics-based models (GNM & ENM), would describe, based on the evidence of agreements of rmsd profiles of 64 NMR-solved protein ensembles and that predicted by GNM and ENM [12].

For a representative set of 64 nonhomologous proteins, each containing a structure solved by NMR and X-ray crystallography, we analyzed the variations in atomic coordinates between NMR models, the temperature (B) factors measured by X-ray crystallography, and the fluctuation dynamics predicted by the Gaussian network model (GNM). The NMR and X-ray data exhibited a correlation of 0.49. The GNM results, on the other hand, yielded a correlation of 0.59 with X-ray data and a distinctively better correlation (0.75) with NMR data. The higher correlation between GNM and NMR data, compared to that between GNM and X-ray B factors, is shown to arise from the differences in the spectrum of modes accessible in solution and in the crystal environment. Mainly, large-amplitude motions sampled in solution are restricted, if not inaccessible, in the crystalline environment of X-ray [12]. Combined GNM and NMR analysis emerges as a useful tool for assessing protein dynamics.

Mechanochemical Activity of Enzymes (with Drs Ivet Bahar) (published [1])
To understand the coupling between collective dynamics and catalytic activity of the enzyme, a systematic analysis of 98 non-homologous enzymes of different EC classes were analyzed using GNM (11, 18), fluctuation patterns near the known catalytic residues (experimentally known) are studied (18). In more than 70% of the examined enzymes, the hinge minima predicted by the GNM are found to be co-localized with the catalytic sites (1). Low rotational mobility (<7%) was observed for the catalytic residues consistent with the functional necessity of enzymes to achieve precise mechanochemical activities (1). These studies agree with previous studies that hinge flexibility is an important mechanism in determining protein conformation and facilitating ligand binding (19), in mediating allosteric effects (21) or fine-tuning enzyme functions (5, 11, 20).
This work has been cited 61 times according to google since published in 2005 June.

Chemistry-free model for Characterizing/Predicting Key Residues that Control Enzymatic Functions (with Drs Ivet Bahar and Akio Kitao) (ongoing [p1])
A de novo algorithm, COnformational-Mobility-based Prediction of enzyme ACTive sites (COMPACT) [p2], was developed to further test the idea that 'enzyme active sites co-localized with the dynamic hinge centers' [1]. *COMPACT*, a protein active site predictor based on coarse-grained structure (shape of the molecule) and dynamic information, has scored a selectivity rate of 0.82 and a moderate to low specificity rate 0.37 over 18 enzymes in different enzyme classes. *COMPACT* basically narrows down potential candidates of catalytic residues, hinge minima, predicted from GNM, to spatially clustered ones with relatively high solvent accessibilities for its position in the protein. This allows the prediction for biomolecules' functional sites without even knowing the amino-acid sequence and atomic coordinates. We repeated a similar procedure on X-ray *electron density maps of enzymes* and were able to locate the position of the functional sites within the electron cloud with a sensitivity and specificity of 0.73 and 0.33 respectively. The restrained dynamics can be of an important active site design properties [20] and the prominent geometrical shape of enzyme is highly related to its active site location. The online implementation of this prediction algorithm can be found in *o*GNM [10] and *Dynomics*.

Intrinsic Dynamic Domains reveal dynamic requirements of protein-protein docking orientation. (with Dr Shun Sakuraba) (ongoing/submitted [15])

A new *parameter-free approach* is proposed to define dynamic domains according to biomolecules' intrinsic collective dynamics. A domain interface plane, defining the separation of domains, and bending axes, about which the domains bend, are as well defined in a parameter-free manner.

The algorithm involves reduction of 3N by 3N inverse Hessian matrix into an N by N covariance matrix that is to be diagonalized to give eigenvectors which define dynamic domains. The so-defined domains are termed 'Intrinsic Dynamic Domains' (IDDs).

As IDDs are examined on all the 8 'difficult' cases in a standard protein-protein docking benchmark, it is found that the interface planes, defined using Linear Discriminant Analysis (LDA) approach, of big proteins always dissect through their small binding partner. In addition, the bending axes of docked protein pairs tend to be perpendicular to each other, which minimize the mutual dynamical interferences of a docking event. A docking scheme that "seeks the tranquility in the move" is proposed to explain dynamics-refrained docking behavior.

Model with Energy Penalty on Inter-residue Rotation gives superior description of crystallographic temperature factors and protein conformational changes. (by myself) (ongoing/submitted [16])

Gaussian Network Model (GNM) is demonstrated NOT rotation invariant for its potential that penalizes the rigid-body rotation. As a result, rigid-body rotation is found to blend in GNM modes, especially in the most collective a few. **A new model is proposed** here to remove such external contributions when maintaining the mode motions that result from a potential that penalizes inter-residue rotation. The new model is shown to well describe crystallographic temperature factors and protein open↔close transition.

***i*GNM Database (with Drs Ivet Bahar et al) (published [2])**

Structural knowledge of a protein is not sufficient for understanding and controlling its function. We have systematically characterized protein dynamics by designing and creating a database of protein functional motions based on the Gaussian Network Model. This database, called *i*GNM (http://ignm.ccbb.pitt.edu) [2], contains the computed GNM dynamics for 20,058 structures from the Protein Data Bank (PDB), and generated information on the equilibrium dynamics at the level of individual residues. *i*GNM is a database of predicted mobilities for all PDB structures (deposited prior to September 1st, 2003), ranging from small enzymes to large complexes and assemblies in a unified framework. The results are stored on a web-based system, and configured so as to permit users to visualize or download the results through a standard web browser using a simple search engine (hence the name, *i*nternet *GNM*). Static and animated images for describing the conformational mobility of proteins over a broad range of normal modes are accessible, along with an online calculation engine available for newly deposited structures. The current version of *i*GNM consists of three modules: Database Engine [2], GNM Computations Engine [10], and Visualization Engine [7]. Since Nov. 2005, we have an updated *i*GNM version (http://ignmtest.ccbb.pitt.edu) for GNm results of all the proteins in PDB (updated weekly). We also parsed the PDB files more carefully and

advanced our search engine to locate GNM data from proteins bearing certain desired features (i.e. size of the protein, X-ray or NMR proteins, resolutions etc.). We also report the results about the correlation between predicted and experimentally observed thermal fluctuations as well as the strength of the spring constant used in the GNM computation.

***o*GNM Web Server (with Drs Ivet Bahar et al) (published [10])**
The web-based system, *o*GNM [10], enables users to calculate online the shape and dispersion of normal modes of motions for proteins, oligonucleotides and user-compiled complexes or monomers using GNM. Computations with new engines, including the implementation of a new algorithm, *PowerB* [10], computing the theoretical thermal fluctuations, are 5-6 order of magnitude faster than those using conventional normal mode analyses. Our implementation also provide a handle to rapid assessments of dynamics of DNA/RNA/Protein complexes by GNM using EN representation that each amino acid and nucleotide is represented by 1 and 3 nodes, respectively, connected by harmonic springs. The rapid computation has facilitated the assessment of thermal fluctuations for a set of 1250 non-homologous proteins. Results have shown a broad range [7.3-15Å] of inter-residue distances can be applied to GNM to get equivalent fluctuations and the correlations with X-ray temperature factors improve with increasing crystallization temperatures.
*o*GNM can be accessed at http://ignm.ccbb.pitt.edu/GNM_Online_Calculation.htm. On the other hand, the anisotropic fluctuations [14] (based on elastic network model) of biomolecules can be obtained and assessed in another webserver of ours, ANM webserver [10a], at http://ignmtest.ccbb.pitt.edu/cgi-bin/anm/anm1.cgi.

The *i*GNM/*o*GNM database/webserver [1,10] as well as ANM webserver [10a] have been made links available in PDB at http://www.pdb.org/pdb/static.do?p=software/software_links/modeling_and_simulation.html ; *i*GNM in fact has been the first dynamics database commensurate with PDB in size. The emerged *Dynomics* project interlinks the existing resources of ours with results from atomistic-NMA based on semi-empirical force fields (*Promode* collaborated with Dr. Wako) and experimentally characterized (mainly by X-ray crystallography) conformational changes (*Dyndom*; collaborated with Dr Hayward).

Study of Polymerase Q (Collaboration with Dr Richard Wood et al.) (published [6])
The basic question we had for this study was why PolQ, a DNA polymerase involved in DNA repair process, has a low fidelity to bypass DNA damages. Given the known PolQ sequence at hand, we performed homology (comparative) modeling to obtain its 3D structure and found an insertion loop docked perfectly in the DNA minor groove despite its flexible nature demonstrated by normal mode analysis (6). We therefore inferred the 'culprit' that influences the fidelity of PolQ to be the insertion 1 loop we identified in this work. This work was published on EMBO journal.

The rational design of enzymes (Collaboration with Dr Alan Russell et al.) (published [3,4])
We studied the cause that attributes to the enzyme activity loss in the immobilized

sarcosine oxidase. A few lysine residues not only bearing their structural and dynamic significance but also possessing abnormal pKa shift, inferred from the software UHBD, were identified to involve in this process (4). The same techniques were also used to rationalize how a cysteine residue, locating far away from the enzyme active site, can possibly interfere with the catalytic ability of creatine amidinohydrolase upon the contamination of silver ions (3) in a biosensor comprised of above-described two enzymes.

Relation-mappings of protein dynamics between experimental and theoretical approaches to reveal the time scales of the dynamic events (with Dr Akio Kitao)(ongoing)

No matter for the "errors/uncertainties" in protein structure determination or equilibrium fluctuations of atoms in proteins at a given temperature, there is a characteristic time to describe the spatial span of such "uncertainties" or fluctuations. By decomposing time-average fluctuations into constituents of frequency components (normal modes/PCA modes), the theoretical methods allow us to subtract or add up a range of these components to obtain the best agreement with experimental observables so as to understand the timescales of our interests, either the experimental observables themselves or fluctuation variance characterized by a given theoretical means.

We substantiated the idea in a comparison of time scales of order parameters of ubiquitin characterized experimentally from NMR relaxation data with that predicted by two theoretical approaches, namely MD simulations and Elastic Network Model using Lipari-Szabo formalism.

$$S_{ij}^2 = \frac{3}{2}(\langle x_{ij}^2\rangle^2 + \langle y_{ij}^2\rangle^2 + \langle z_{ij}^2\rangle^2 + 2\langle x_{ij}y_{ij}\rangle^2 + 2\langle x_{ij}z_{ij}\rangle^2 + 2\langle y_{ij}z_{ij}\rangle^2) - \frac{1}{2}$$

We first performed a long 120ns MD simulation and then obtain the PCA modes of its trajectory. Instead of getting N-H bond order parameters (OP) directly from all the snapshots using above equation, we first obtain a new set of snapshots that is generated from subsequent removal of contributions of low modes to positional deviations from the mean and then obtain OP for the resulting set of snapshots. We identify how many modes needed to be deleted to have the best experimental and theoretical agreement of OP and find the characteristic time from the correlation function of PC-coordinates of that very mode. We successfully obtain the characteristic time is slightly less than 1 ns for OP of ubiquitin, which is the generally accepted time scale for order parameters. We can then adopt similar approach to gauge the time scale of elastic network model from first identifying the mode until which slow modes are subsequently deleted to give the best agreement with experimental data. This technique is expected to explore the 'real' time scale of CG-MD as well as other experimental observables such as isotropic or anisotropic temperature factors.

Profile of transmembrane dipolar potential is a function of membrane tension (with Dr M. Chachisvilis)

It has been known that physical forces applied to tissues, organs or blood vessels could trigger their down stream biochemical signal propagation. Our hypothesis here is that the mechanical forces (such as applied a tension on lipid membranes) could cause membrane morphology as well as dipole potential changes, which in turn cause the conformational changes of membrane proteins. We used MD simulations while maintaining constant membrane tension to investigate the dipole potential changes across the membrane. The dipole potential profiles predicted from simulation well agree with that suggested from experimental data. The next step would be to insert a membrane protein in the membrane and see how the dipole potential changes would influence the conformation of membrane protein.

Pre-steady state enzyme kinetics studies on octaprenyl pyrophosphate synthase and undecaprenyl pyrophosphate synthase (with Dr P.H. Liang)(published [8,9])

The study is about cell wall sugar elongation process catalyzed by two enzymes, octaprenyl pyrophosphate synthase and undecaprenyl pyrophosphate synthase (potential anti-bacteria target). We were able to determine the kinetic of intermediate steps of this elongation process. In Liang's lab, the techniques acquired include PCR, site-directed mutagenesis, sequencing technique, isotope assays used in kinetic study, concepts of pre-steady-state techniques including stop-flow operation.

Reference

Published Journal Articles:

1. **Yang, L.-W.** and Bahar, I. (2005) Coupling between catalytic site and collective dynamics: A requirement for mechanochemical activity of enzymes. *Structure*, 13, 893.
2. **Yang, L.-W.**, Liu, X., Jursa, C.J., Holliman, M., Rader, A.J., Karimi, H.A., Bahar, I. (2005) iGNM: a database of protein functional motions based on Gaussian Network Model. http://ignm.ccbb.pitt.edu/ *Bioinformatics*, 21, 2978-2987.
3. Berberich,J., **Yang, L.-W.**, Bahar,I., Russell,A.J. (2005) Analysis of the impact of silver ions on creatine amidinohydrolase. *ACTA Biomateriala* 1, 183-191.
4. Berberich,J., **Yang, L.-W.**, Madura, J., Bahar,I., Russell,A.J. (2005) A stable three enzyme creatinine biosensor impact of structure, function and environment on PEGylated and immobilized sarcosine oxidase. *ACTA Biomateriala* 1, 173-181.
5. Chennubhotla, C., Rader, A.J., **Yang, L.-W.**, Bahar, I. (2005) Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies. *Phys. Biol.*
6. Seki, M., Masutani, C., **Yang, L.-W.**, Schuffert, A., Iwai, S., Bahar, I. and Wood, R.D. (2004) High efficiency bypass of DNA damage by a single human DNA polymerase. *EMBO J.* 23, 4484-4494.
7. Liu, X., Karimi, H., **Yang, L.-W.**, Bahar,I. (2004) Protein functional motion query and visualization" IEEE Proceedings, 28th Annual International Computer Software and Applications Conference - (COMPSAC'04)
8. Pan, J.J., Kuo, T.H., Chen, Y.K., **Yang, L.-W.**, Liang P.H. (2002) Insight into the activation mechanism of Escherichia coli octaprenyl pyrophosphate synthase derived from pre-steady-state kinetic analysis. *Biochim Biophys Acta.* 1594, 64-73.
9. Pan, J.J., **Yang, L.-W.**, Liang, P.H. (2000) Effect of site-directed mutagenesis of the conserved aspartate and glutamate on E. coli undecaprenyl pyrophosphate synthase catalysis. *Biochemistry.* 39, 13856-13861.
10. **Lee-Wei Yang**, A.J. Rader, Xiong Liu, Cristopher Jon Jursa, Shann Ching Chen, Hassan Karimi and Ivet Bahar (2006) oGNM: A protein dynamics online calculation engine using the Gaussian Network Model *Nucleic Acids Res,* 34**,** W24-31 http://ignm.ccbb.pitt.edu/GNM_Online_Calculation.htm

10a. Eyal, E., **Yang, L.-W.,** Bahar, I. (2006) Anisotropic Network Model: systematic evaluation and a new web interface. http://ignmtest.ccbb.pitt.edu/cgi-bin/anm/anm1.cgi *Bioinformatics,* 22, 2619-2627

11. Rader, A.J., Chennubhotla, C., **Yang, L.-W.**, Bahar, I. (2005) "The Gaussian Network Model: Theory and Applications." Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems. *CRC Press.*
12. **Yang, L.-W.,** Eyal, E., Chennubhotla, C., Jee, J., Gronenborn, AM., Bahar, I. (2007) Insights into Equilibrium Dynamics of Proteins from Comparison of NMR and X-ray Data with Computational Predictions. *Structure*, 15, 741-749
13. **Yang, L.-W**., Eyal, E., Bahar, I. and Kitao, A. (2009) Principal Component Analysis of Native Ensembles of Biomolecular Structures (PCA_NEST): Insights into Functional Dynamics. *Bioinformatics*, 25, 606-614 http://ignm.ccbb.pitt.edu/oPCA_Online.htm
14. **Yang, L-W.**, Kitao, A. and Gō, N. (2010) Primary Responses of Myoglobin Examined by Linear Response Theories. (Submitted; submission proof)

15. **Yang, L-W.** and Sakuraba, S. (2010) Dynamic predispositions of protein-protein docking orientation revealed by Intrinsic Dynamic Domains. (Submitted; submission proof)
16. **Yang, L-W.** (2010) Model with Energy Penalty on Inter-residue Rotation gives superior description of crystallographic temperature factors and protein conformational changes. (Submitted; submission proof)

Manuscripts in preparation:

p1. **Yang, L.-W**., Kitao, A. and Bahar, I. (2011) Chemistry-free models reveal the geometrical predisposition of the location of enzyme active sites. (to-be-submitted)

Publications (by others) dearly related to my research

17. Atilgan, AR, Durrell, SR, Jernigan, RL, Demirel, MC, Keskin, O. and Bahar, I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophys. J. 80, 505-515, 2001.
18. Bahar, I., Atilgan A.R., and Erman, B. Direct evaluation of thermal fluctuations in protein using a single parameter harmonic potential. *Fold. Des.* 2, 173-81 (1997); Haliloglu, T. Bahar, I. and Erman. B. *Phys. Rev. Lett.* 79, 3090-3093 (1997).
19. Bahar, I., Atilgan, A.R., Demirel, M.C. and Erman, B. Vibrational Dynamics of Folded Proteins: Significance of Slow and Fast Motions in Relation to Function and Stability. *Phys. Rev. Lett.* 80, 2733-2736 (1998).
20. Gutteridge, A., Thornton, J.M. and Bartlett,G. Using a neural network and spatial clustering to predict the location of active sites in enzymes. *Biochemistry* 37, 11940-11948 (2003).
21. Xu, C., Tobi, D. And Bahar,I. Allosteric changes in protein structure computed by a simple mechanical model: Hemoglobin T ↔ R2 transition. *J. Mol. Biol.* 333, 153-168 (2003).
22. S. Hayward and N. Go, Annu. *Rev. Phys. Chem.* 46, 223 (1995)
23. Ikeguchi, M., Ueno, J., Sato, M. and Kidera, A. Protein structural change upon ligand binding: linear response theory. *Phys Rev Lett* 94, 078102-1-078102-4 (2005)
24. Dror Tobi and Ivet Bahar (2005) *Proc Natl Acad Sci* **102**, 18908
25. Chandrasekhar, S. (1943) Stochastic Problems in Physics and Astronomy. *Rev Mod Phys*, **15**, 1-89