Journal of Engineering and Natural Sciences
Mühendislik ve Fen Bilimleri Dergisi

Sigma 27
161-170,
2009

Invited Review Paper / Çağrılı Derleme Makalesi

PROTEIN DYNAMICS

Pemra ÖZBEK, Türkan HALİLOĞLU*

Polymer Research Center and Chemical Engineering Department, Boğaziçi University, Bebek-ISTANBUL

Received/Geliş: 24.08.2009

ABSTRACT

Proteins are building units of genomes. There exist a correlation between sequence, structure and function of proteins, and function is closely related to structural dynamics. To understand how genomes and proteomes give rise to biological functions, it is necessary to go from the three to the fourth dimensional pictures of proteins; that is their conformational space. Proteins are not static but dynamic entities and they undergo conformational changes described by their multidimensional energy landscape for their function. The dynamic nature of proteins depends on them being flexible. The observed conformational changes may involve both local motions involving a few residues and global cooperative motions of several residues. The information about the flexibility and dynamics of proteins could partially be inferred through experimental means, yet computational approached in recent years has complemented experimental studies and contributed to the understanding on how proteins are in action. A few case studies are presented here demonstrating the use of computational techniques.
Keywords: Protein dynamics, conformation, molecular simulations, elastic network models.

PROTEİN DİNAMİĞİ

ÖZET

Proteinler genomların önemli yapı taşlarındandır. Proteinlerde dizi-yapı-işlev paradigmasının anlaşılması genomun işlevinin anlaşılmasında önemli bir aşamadır. Proteinlerin üç boyutlu yapılarından dört boyutlu yapılarına bu proteinlerin konformasyonal uzayının elde edilmesi ile gidilebilir ve bu proteinin işlevinin anlaşılması için önemlidir. Proteinler statik konumda olmayıp dinamik haldedirler, işlevlerini yapmak üzere konformasyon değişiklikleri yaparlar. Dinamik yapıları esnek olmalarından kaynaklanır. Gözlenen konformasyonal değişiklikler yerel ve yalnızca birkaç aminoasiti içeren hareketlere sebebiyet verebildiği gibi birçok aminoasitin uzayda konumunu değiştirmesini gerektiren global -kolektif- hareketler anlamına da gelebilir. Proteinlerin konformasyonal esnekliği ve dinamiği ile ilgili şu anda var olan bilgi kısmen deneysel çalışmalardan gelmekte olup, son zamanlarda geliştirilen hesapsal yaklaşımların da katkılarıyla proteinlerin davranışlarını anlamak mümkün olabilmektedir. Burada hesapsal yöntemler ile protein dinamiğinin çalışıldığı bir kaç örnek çalışma verilmiştir.
Anahtar Sözcükler: Protein dinamiği, konformasyon, moleküler simulasyon, elastik ağ yapı modelleri.

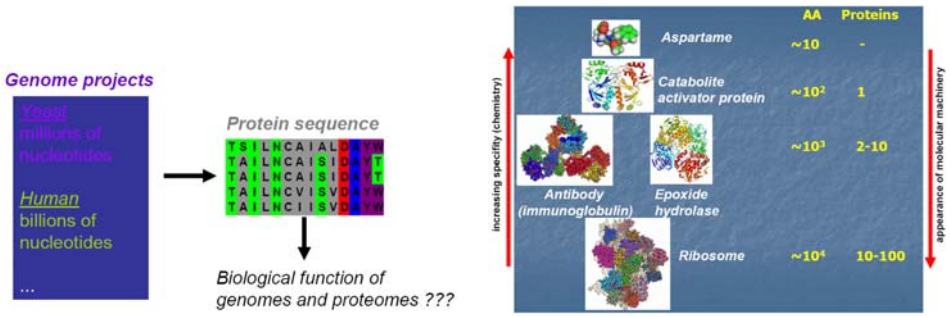
Proteins Are In Motion

In living organisms, almost all biological processes require protein dynamics. Proteins are in continuous motion through sampling ensemble of conformations, interacting with other proteins,

*Sorumlu Yazar/Corresponding Author: e-mail / e-ileti: halilogt@boun.edu.tr, tel: (212) 359 70 03

or other biological molecules, as they function in a particular biological process. To this end, it is of significant interest to go from the structures to the ensemble of conformations, where their dynamic personalities are described. [1]

How is the connection between dynamics and function? With the progress of genomics projects, a large number of nucleotide and amino acid sequences have become available which are collected in the related databanks (Fig. 1.a). Along, thousands of low- to high-resolution protein structures are obtained with the efforts in the structural genomics field. Nevertheless, the knowledge of sequence and structure is not enough to understand how genomes and proteomes work.



(a) **(b)**

Figure 1. a) Amino acid sequences for biological function b) Hierarchical organization of biological molecules from amino acids to molecular machines.

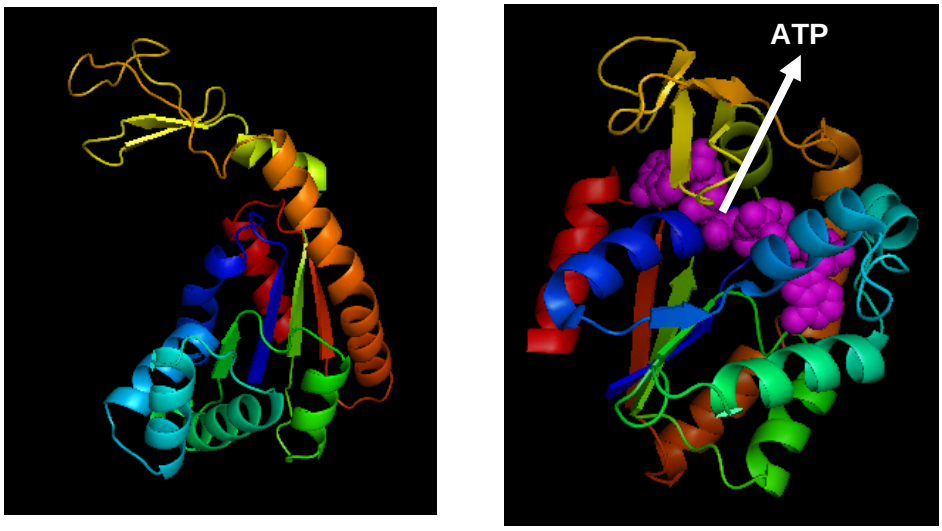
Life could be described with a complex pyramid architecture [2], where bottom level (level 1) consists of cell's functional organization; genome, transcriptome, proteome, metabolome, which are the constituents of protein-protein interaction networks and metabolic pathways (level 2) that take part in the formation of the functional modules (level 3) making up the organism (level 4). Yet, every stage of the pyramid has varying scales of structural organization (Fig. 1.b). At the bottom level, for example; we see proteins, made up of amino acids, associating (binding) to form protein complexes at a scale with high specificity to a scale where we observe the appearance of molecular machinery. Bottom level is the level where we observe the complexity along the diversity the most.

At the bottom level, sequence-structure-function paradigm is of significant interest with the ultimate aim of understanding the protein function. The sequence of amino acids determine the three dimensional structure of the protein (protein folding) [3, 4]. In the determination of the structure, X-ray Crystallography [5] and Nuclear Magnetic Resonance (NMR) [6] are the leading experimental methods. Additionally, electron microscope (Cryo-EM) [7] has increasingly been used to determine the three dimensional structure of large protein complex structures (assemblies). On the other hand, computational methods with the efficient use of bioinformatics tools have gained significant success in the prediction of protein structures from high- to low-resolution [8-10] in recent years. Yet, determination of the function of a protein structure in the expression of genetic information remains as one of the most challenging areas of research.

Sequence-Structure-Function Paradigm: Function and Dynamics

In the sequence-structure-function paradigm, a close relationship between the function and the structure's dynamic properties appears. Almost all biological processes that require action are through means of protein dynamics. Protein structures are not static and they change their conformation in space to achieve their function. Besides of the proteins whose main function is

mobility, dynamics is also important for transport proteins, signaling proteins, immune proteins and enzymes. For many enzymes, conformational changes are important for binding to a substrate in addition to controlling the rate of a catalytic activity. For example, adenylate kinase (phosphotransferase enzyme), which is an important enzyme controlling the balance of ATP in prokaryotic cells, undergoes large conformational changes during catalytic cycle as can be seen in Fig. 2, where both open (unbound) and close (bound) conformations are shown.



(a) **(b)**

Figure 2. Adenylate Kinase (a) open state without bound inhibitor (PDB entry 4ake) (b) closed state with bound inhibitor ATP (PDB entry 1ake)

Proteins are dynamic in their nature, and their function depends on them being flexible. [11] Being dynamic, proteins alter their conformation for their intrinsic motions that are controlled by the covalent and non-covalent forces of the atoms in their structures. The conformational space that a protein can travel can be described by an energy landscape (Fig. 3), in which different conformations are taken based on their energies, and interconversion rates are dependent on the energy barriers between states. The time scales of the dynamics are associted with these barriers. Protein dynamics is dominated by harmonic motions at low temperatures; the contribution of non-harmonic motions may become relevant as the temperature increases. The fastest conformational rearrangements, which are mostly localized motions, involves a few aminoacids and can take place in picosecond time scales. Yet, the collective motions take place in nanosecond time scales, while large-scale events (such as folding, molecular binding processes) occur more slowly in milliseconds or even seconds, involving several aminoacids of the structure. The observed dynamics is an interplay of motions at frequencies from picoseconds to milliseconds or even seconds. [12]

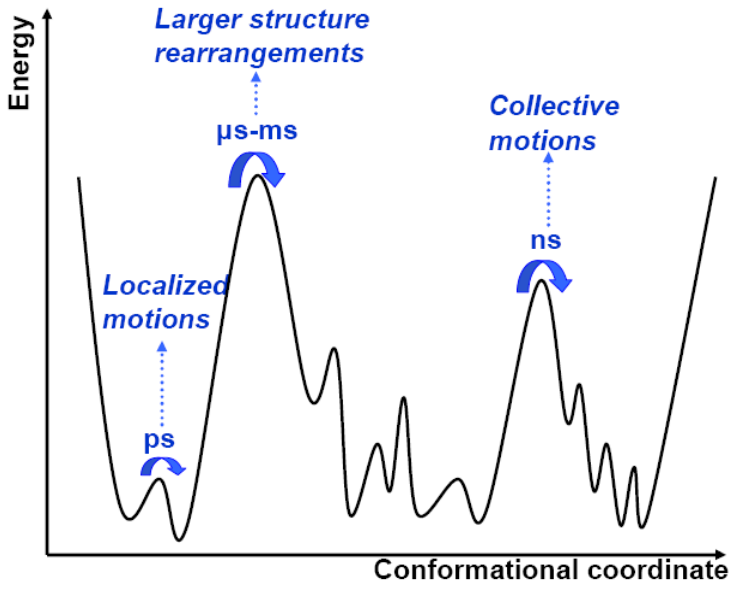


Figure 3. Energy landscape; energy vs. conformational coordinate

The landscape can be modulated, meaning that the relative populations of conformational states change. [13] This can happen, for example, by interactions with other proteins (physical interactions or covalent modifications), or by variations in the environment such as temperature change. The reasons that modulate the energy landscape at the same time might disrupt or control protein function. Today's efforts in structure-based drug design could have more opportunites by considering the entire-free energy landscape, instead of targeting a static active site. With this, future efforts could lead to allosterically active drugs or a drug that acts by trapping proteins in an inactive state. [13] Such drugs can be designed to inhibit enzyme activity or to inhibit an interaction on a pathway. This puts forward that the alternate functions described by the minor conformers may become more active with the mutations that leads to the shift in the conformational equilibrium to favor these conformers. [14] This implies that the conformational diversity of proteins makes them also evolvable and thus could provide the grounds for evolutionary adaptations as recently discussed. [14]

Experimental and Computational Approaches

The current knowledge on protein conformational flexibility is derived from X-ray Crystallography and Nuclear Magnetic Resonance (NMR). [5, 6] NMR experiments used to be limited to small, soluble proteins. Recent developments in the spectrometer technology (high magnetic fields and cryoprobes) and new NMR pulse sequences enabled the study of proteins even up to the size of a ribosome. [15-21] To understand proteins in action, the changes of the coordinates in time must be added to the snapshots of proteins. Here, the major bottleneck is that experimentally it is not possible to watch individual atoms moving within a protein, i.e, to have the movie of the macromolecular structure. Instead, complex biophysical methods are needed to be designed to measure the physical properties from which the dynamics can be deduced. [1] For example; fluorescence resonance energy transfer (FRET) method enables the study of the function of a single molecule in real time by fluorescent dyes using distance dependent interaction

as a measurement parameter. [22] This approach yields a continuous observation of a distance that is reminiscent of a functional conformational change, as recently presented [23] for the mechanism of catalysis during F_0F_1-ATP synthesis. With all, the knowledge on the transition process among one state to another is still quite limited. Today, on the other hand, it is possible to display the behavior of the proteins at various time scales using computational approaches. In these methods, the timescale, amplitude and the directions of motion play an important role. Molecular Dynamics (MD) simulation is the most widely used technique to describe the protein motions in atomic scale in real time. But still, protein dynamics on the larger time scales, such as microsecond-to-millisecond, is currently out of reach for molecular-dynamics simulations. To achieve longer time scales, various approaches with simplified force fields have been developed. Normal mode analysis (NMA), studying harmonic potential wells by analytical means, has gained notable success in the area of dynamics of biomolecules. [24 -26] Simpler models based on polymer network mechanics have been introduced to apply NMA such as Gaussian network model (GNM), which is an elastic network (EN) model introduced at residue level. [27, 28] Although being a simple approach, GNM, and Anisotropic Network Model (ANM) [29], or similar coarse-grained EN models have found widespread use in the study of protein dynamics. FIRST (floppy inclusion and rigid substructure topography) [30], FRODA (framework rigidity optimized dynamic algorithm) [31], and Gō models [32] are the other methods that has been used in this area. Additionally, small scale network models and potentials are also used to predict molecular behaviour such as: Brownian Dynamics (BD) and Monte Carlo (MC) Simulations. [33]

Here, we provide below some case studies where the dynamics is inferred for the function by various computational tools.

Case Studies

Calmodulin

Calmodulin is a calcium binding protein affecting many different cellular functions. It undergoes very large conformational changes upon binding to calcium, which enables it to bind to other specific proteins. The previous NMR studies have shown that a hinge region from Asp 78 to Ser 81 exists as being highly flexible. [34, 35]

Hingeprot [36], having been developed by making use of GNM [27, 28] and ANM [29], enables the prediction of rigid protein parts and the flexible hinge regions connecting them in the native topology of protein chains. These hinge regions can help to understand functional mechanisms of macromolecular structures and assemblies. For Calmodulin, Hingeprot results show (Fig. 4) that the first slowest mode gives a hinge region at Asp 80/Ser 81 residues on the long helix, splitting the structure into two rigid parts (domains) as (Ala 1-Asp 80) and (Ser 81-Lys 148) [36], which is in agreement with NMR studies. [34, 35]

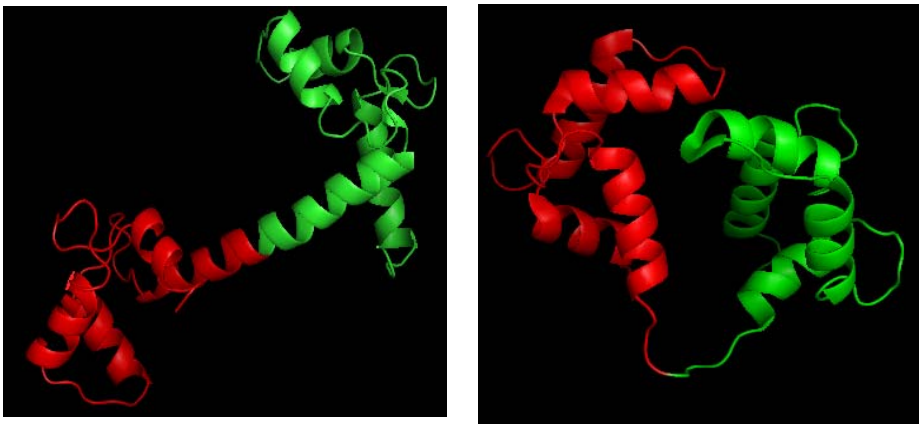
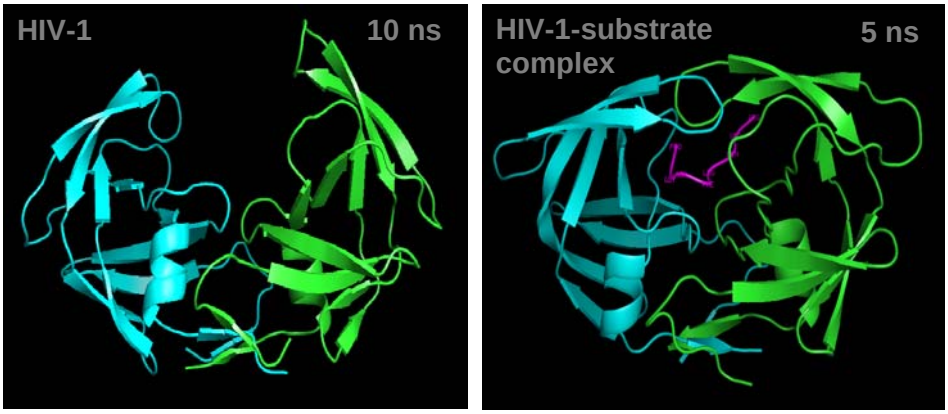


Figure 4. Conformations of Calmodulin protein in unbound (PDB entry 4cln) and bound (PDB entry 2bbm) states. The slowest first mode by Hingeprot [36] divides the protein into two domains: (Ala 1-Asp 80) and (Ser 81-Lys 148), where the hinge residue at Asp 80/Ser 81 can be observed.

HIV-1 Protease

HIV-1 protease is an essential enzyme for the replication of HIV (Human Immunodeficiency Virus), the virus that causes AIDS. This protease is thus a primary target for the development of drug design against AIDS. The understanding of the dynamics of unbound and substrate bound HIV-1 protease is essential to guide the design of next generation HIV drugs. Figure 5 displays a snapshot from MD simulations of unbound protease and substrate-bound protease. The analysis of MD simulation with the predictions of the GNM contributes to the understanding of the protease dynamics comprehensively, pointing to the fluctuations and the key residues that are essential for the protease function and stability. [37]



(a) **(b)**

Figure 5. Cartoon diagrams of (a) unliganded and (b) liganded HIV-1 protease crystal structure. The two monomers are in cyan green, the substrate in (b) is in magenta.

Potassium Channels

Potassium channels, found in all types of cells, are involved in key biological processes. Taking part in the generation and propagation of nerve impulses, they switch between open and closed conformations. They control the ion transfer through the channel pore and biological cell functions, thus understanding the switch mechanism and the structural characteristics of the transition between conformations are important. The proper function of the channel in the cell depends on the ability of the structure to undergo conformational transitions. Controlling the open/closed states of the pore in response to external signals, such as ligand binding or changes in the membrane potential is achieved by these conformational transitions. [38] The comparison of the structures of the KcsA channel in its closed and open conformations (Fig. 6) show that the main change is identified mainly with a bend of the inner helices of the four subunits. Using the EN models, namely GNM [27, 28] and ANM [29], it has been possible to show that there exists a network of energetically and dynamically coupled residues in the complex structure of four subunits. The interactions between the subunits are essential for the coupling between the two gates of the channel. [38] Figure 6 displays the fluctuations (a, c) around the crystal structure (b) by ANM, which reflect the opening (a) and closing (c) behavior of the channel.

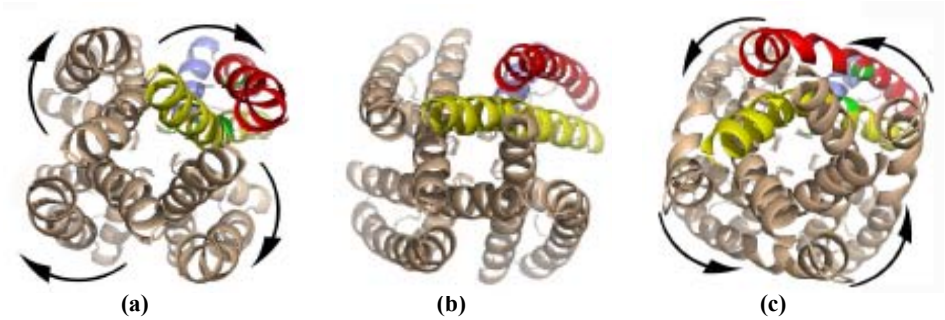


Figure 6. Predicted conformations of KcsA by elastic network models [37]

Peptide Membrane Interactions

Using Monte Carlo (MC) methods, peptide-membrane systems can be studied via reduced representation of the system when compared with the molecular dynamics (MD) simulations which use all-atom representation. By reduced representation, comprehensive sampling of peptide conformations and locations in the membrane can be achieved in an accelerated manner. [39] This can be seen from the snapshots taken from the MC simulations (Fig. 7) of the hydrophobic peptide, M2δ, with a model membrane. The peptide, having started with an extended conformation in the aqueous phase, associates with the membrane surface by assuming a helical conformation and gets into the membrane eventually. The calculations suggest that there are two main configurational states for the peptide, transmembrane configuration being with higher probability. The calculated values of free energies of transfer and the observed conformational changes are in agreement with the previous experimental studies, confirming the accuracy of this method in the prediction of three-dimensional structure of Trans membrane (TM) peptides. [39]

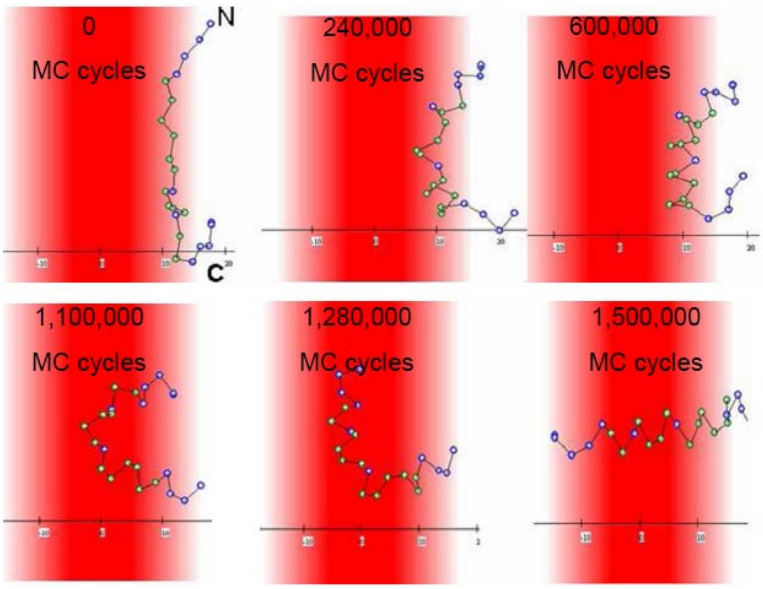


Figure 7. Peptide membrane interaction. Folding of M2δ (AchR) on the membrane surface, starting from an extended conformation

Today, with efforts to search for a relation between computational approaches, bioinformatics tools and experimental methods, a scientific progress in solving the sequence-structure-function paradigm has been achieved; nevertheless there is still a long way to go.

Acknowledgements

P. O.'s research is supported by TUBITAK (2218, Research Fellowship).

REFERENCES / KAYNAKLAR

[1] Henzler-Wildman, K. and Kern, D., Dynamic personalities of proteins, Nature, 13:450 (7172):964-72, 2007.

[2] Oltvai, Z. N. and Barabasi, A.-L., Life's Complexity Pyramid, Science, 298, 763-764, 2002.

[3] Baker, D. and Lim, W.A., From folding towards function, Current Opinion in Structural Biology, 12, 11-13, 2002.

[4] Clarke, J. and Schreiber, J., Folding and binding - new technologies and new perspectives, Current Opinion in Structural Biology, 13, 71-74, 2003.

[5] Azároff, L. V., Kaplow R., Kato N., Weiss R. J., Wilson A. J. C. and Young R.A., X-ray diffraction. McGraw-Hill, 1974

[6] Martin, G. E. and Zekter, A. S., Two-Dimensional NMR Methods for Establishing Molecular Connectivity. VCH Publishers, New York, p. 59, 1988.

[7] Frank, J., Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Oxford University Press, New York, 2006.

[8] Zhang, Y., Progress and challenges in protein structure prediction, Current Opinion in Structural Biology 18 (3): 342-348, 2008

[9] Zhang, Y. and Skolnick, J., The protein structure prediction problem could be solved using the current PDB library, Proc Natl Acad Sci USA 102 (4): 1029-1034, 2005.
[10] Bradley, P., Misura, K. M. S. and Baker, D., Toward High-Resolution de Novo Structure Prediction for Small Proteins, Science, 309 (5742): 1868 – 1871, 2005.
[11] Teilum K., Olsen J. G. and Kragelund B. B., Functional aspects of protein flexibility, Cellular and Molecular Life Sciences, 66:14, 2231-2247, 2009.
[12] Smock, R. G. and Gierasch, L. M., Sending signals dynamically. Science, 10;324(5924):198-203, 2009.
[13] Vinson V. V., Proteins in Motion, Science, 324 (5924): 197, 2009.
[14] Tokuriki, N. and Tawfik, D. S., Protein Dynamism and Evolvability, Science 324, 5924, 203 – 207, 2009.
[15] Pervushin, K., Riek, R., Wider, G. and Wuthrich, K., Attenuated T-2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc. Natl Acad. Sci. USA 94, 12366–12371, 1997.
[16] Sprangers, R., Gribun, A., Hwang, P. M., Houry, W. A. and Kay, L. E., Quantitative NMR spectroscopy of supramolecular complexes: dynamic side pores in ClpP are important for product release. Proc. Natl Acad. Sci. USA 102, 16678–16683, 2005.
[17] Palmer, A. G., Grey, M. J. and Wang, C. Y., Solution NMR spin relaxation methods for characterizing chemical exchange in high-molecular-weight systems. Methods Enzymol. 394, 430–465, 2005.
[18] Tugarinov, V. and Kay, L. E., Quantitative C-13 and H-2 NMR relaxation studies of the 723- residue enzyme malate synthase G reveal a dynamic binding interface. Biochemistry 44, 15970–15977, 2005.
[19] Sprangers, R. and Kay, L. E., Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature, 445, 618–622, 2007.
[20] Horst, R. *et al.* Direct NMR observation of a substrate protein bound to the chaperonin GroEL. Proc. Natl Acad. Sci. USA 102, 12748–12753, 2005.
[21] Christodoulou, J. *et al.* Heteronuclear NMR investigations of dynamic regions of intact Escherichia coli ribosomes. Proc. Natl Acad. Sci. USA 101, 10949–10954, 2004.
[22] Förster, T., Zwischenmolekulare Energiewanderung und Fluoreszenz, Ann. Physik 437, 55, 1948
[23] Diez, M. *et al.* Proton-powered subunit rotation in single membrane-bound F0F1-ATP synthase. Nature Struct. Mol. Biol. 11, 135–141, 2004
[24] Cui, Q. and Bahar, I., Normal Mode Analysis: Theory and applications to biological and chemical systems, Chapman&Hall/CRC, 2005
[25] Case, D. A., Normal mode analysis of protein dynamics. Curr Opin Struct Biol. 4:285–290, 1994
[26] Kitao, A. and Go, N., Investigating protein dynamics in collective coordinate space. Curr Opin Struct Biol. 9:164–169,1999.
[27] Bahar, I., Atilgan, A. R. and Erman, B., Direct evaluation of thermal fluctuations in proteins using a single parameter harmonic potential, Folding & Design, 2, 173-181, 1997.
[28] Haliloglu, T., Bahar, I. and Erman, B. Gaussian dynamics of folded proteins. Phys. Rev. Lett. 79, 3090–3093, 1997.
[29] Atilgan, A. R., Durell, S. R. , Jernigan, R. L., Demirel, M. C., Keskin, O. and Bahar, I., Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophys J. 80:505–515, 2001
[30] Jacobs, D. J., Rader, A. J., Kuhn, L. A. and Thorpe, M. F., Protein flexibility predictions using graph theory. Proteins 44, 150–165, 2001.

[31] Wells, S., Menor, S., Hespenheide, B. and Thorpe, M. F., Constrained geometric simulation of diffusive motion in proteins. Phys. Biol. 2, S127–S136, 2005.
[32] Scheraga, H. A., Khalili, M. and Liwo, A., Protein-folding dynamics: overview of molecular simulation techniques. Annu. Rev. Phys. Chem. 58, 57–83, 2007.
[33] Leach, A.R., Molecular Modelling 2nd, Prentice Hall, 2001.
[34] Larsson, G., Schleucher, J., Onions, J., Hermann, S., Grundstrom T. and Wijmenga, S.S., A novel target recognition revealed by calmodulin in complex with the basic helix-loop-helix transcription factor SEF2-1/E2-2, Protein Sci. 10, 169-186, 2001.
[35] Elshorst, B., Hennig, M., Forsterling, H., Diener, A., Maurer M., Schulte, P., Schwalbe, H., Griesinger, C., Krebs, J., Schmid, H., Vorherr, T. and Carafoli, E., NMR Solution Structure of a Complex of Calmodulin with a Binding Peptide of the Ca2+ Pump. Biochemistry 38, 12320-32, 1999.
[36] Emekli, U., Schneidman-Duhovny, D., Wolfson, H. J , Nussinov, R. and Haliloglu, T., HingeProt: Automated Prediction of Hinges in Protein Structures. Proteins 70:1219-27, 2008.
[37] Kurt, N., Scott, W. R., Schiffer, C. A. and Haliloglu, T., Cooperative fluctuations of unliganded and substrate-bound HIV-1 protease: a structure-based analysis on a variety of conformations from crystallography and molecular dynamics simulations, 15;51(3): 409-22, 2003.
[38] Haliloglu T. and Ben-Tal N., Cooperative Transition between Open and Closed Conformations in Potassium Channels. PLoS Comput Biol 4(8), 2008.
[39] Kessel, A., Haliloglu, T. and Ben-Tal, N., Interaction of Hydrophobic Peptides with Lipid Bilayers: Monte Carlo Simulations with M2d, Biophysical Journal, 85, 3431-3444, 2003.