Native Proteins as Physical Networks: Energy and Geometry Fluctuations and Their Relation to Function

B. Erman

A correspondence between networks and gels, physical or chemical, and native globular proteins is established by visualizing proteins as physical gels. The emphasis is placed on spatial fluctuations of residues in proteins which are correlated, and it is stated that these correlations determine the function of the proteins. Structural and topological features necessary for a gel to exhibit features of a protein such as to perform a predetermined function are discussed. A simple mathematical model that explains the relationships between correlated fluctuations of residues and the function of the protein is given.

Keywords: fluctuations; fluctuation correlations; gaussian network model; structure-function relations

Introduction

The aim of this paper is to establish a correspondence between networks and gels, physical or chemical, and native globular proteins. The three dimensional structures and functions of native proteins are well documented.[1] Here, we try to show that native proteins can be visualized as physical gels. At the nano-scale, the most important feature of gels is the spatial fluctuations of their junctions. In proteins, the fluctuations of amino acid residues, which we identify with junctions of a physical gel, are correlated, and these correlations determine the function of the proteins. We try to answer the question: What structural and topological features are necessary for a gel to exhibit features of a protein such as to perform a predetermined function as a protein does? In order to answer this question, we use a simple mathematical model that explains the relationships between correlated fluctuations of residues and the function of the protein. Relating the behavior of networks and gels to that of proteins in general promises a new and rigorous field of functional synthetic nano-gels, leading to biomimetics at the molecular level. We reemphasize here that the focus of interest is the fluctuations of junctions but not the macroscopic deformations of gels. Macroscopic deformations such as swelling of responsive gels have been the focus of a wide body of work.[2]

Efforts along forming nano-gels, with emphasis on the fluctuations of junctions, and characterizing the networks in terms of their function do indeed exist in the literature of recent years. Supramolecular gels that are sensitive to external stimuli, referred to as smart gels for sensing are examples as reviewed by Sangeetha and Maitra.[3] An important parameter that determines the topological features of a network is its junction functionality, defined as the number of other junctions neighboring a given junction. In this respect proteins, regarded as physical networks, have large junction functionalities. On a coarse grained approximation, with the chiral carbon as the representative point of a given amino acid, the average functionalities are in the range 8–12 in proteins. Spatially inhomogeneous junction functionalities, or cross-link densities, are requisite for functional gels. Shibayama

Department of Chemical and Biological Engineering, Koc University, Sariyer, 34450 Istanbul Turkey
E-mail: berman@ku.edu.tr

et al.[4] showed that increasing cross-link densities cause cross-link inhomogeneities and result in larger scattering from gels. In proteins, inhomogeneities of local amino acid density are closely related to structure and function. The active site of a protein usually has densely packed amino acids needed for a coherent region that will perform the function. The second requirement for function is the presence of secondary structure which was recognized for polymeric systems by Xu[5] who stated that ''molecular self-assembly can confer well defined secondary structures in a liquid that initiates functions within biological systems''. The third requirement is the presence of active centers resulting from the specific chemical nature of the atoms that confer the desired mechano-chemical activity to the gel. This is at the atomic length scale, but the correlations resulting from these interactions extend over the full material. Anion responsive hydrogels,[6] gels that undergo chemo-mechanical instabilities,[7] the wide class of electro, magneto, pH, and thermoactive gels are conceptual equivalents of proteins, whose functions fall into similar categories, but at different time and length scales.[8–11] Several examples are given by Liu and Urban.[12] The papers cited in this paragraph are only a limited cross-section of works published in the last two years.

In the present paper, we give a detailed discussion of topology, molecular architecture and function relations in native globular proteins, with the expectation that the simple model described here will lead to a molecular level understanding of junction fluctuation correlations in responsive gels. In the first part, we briefly describe native globular proteins, followed by a discussion of function in proteins that result from correlated fluctuations of its amino acids, and the relation fluctuation correlations to molecular topology. The molecular model adopted is a simple statistical mechanical model of a harmonic system that is capable of predicting the fluctuation amplitudes in proteins and the emergence of function from correlated fluctuations. We give several examples, and discuss possible transfer of the knowledge from proteins to responsive artificial gels.

Proteins and the Gaussian Network Model, GNM

Proteins are polypeptide chains of covalently bonded amino acid repeat units called residues. There are 20 different types of amino acids. Their sequence along the chain is called the primary sequence. Thus, a protein can be visualized as a copolymer of twenty different types of monomers. The organization of the repeat units into a covalently bonded polymer chain is referred to as the primary structure. Different portions of the primary sequence take various conformations such as helices, beta strands, loops, turns, chain ends which are referred to as the secondary structure of the chain. The secondary structure units arrange in space to give the three dimensional structure of the native protein.[13] The native states of a protein are marginally stable, and of lowest free energy by necessity. In the native state, i.e. the equilibrium conformation in which the protein performs its function, a protein has a well-defined three-dimensional shape with a density, approximately, of a glassy polymer. The stability of this state is marginal, in which residues that have well defined mean positions are not in a frozen state, and, fluctuations take place with amplitudes that may reach nanometers. Each residue has 6 to 12 spatial neighboring residues. Interactions between a given residue and its spatial neighbors are of non-bonded character, van der Waals, hydrogen bonding or electrostatic. In this respect, each residue acts as a physical cross-link, and a protein chain of n residues acts as a nano-physical-network of m cross-links. A spherical protein of 250 residues has an approximate diameter of 50 Å in the native state. The spatial positions and fluctuations of residues, corresponding to the junctions of the physical network are determined by x-ray crystallography.[1] The relationship between the amplitude of atomic fluctuations and the measured

quantity, B factor is

$$\left\langle (\Delta R_i)^2 \right\rangle = \frac{3}{8\pi^2} B_i \quad (1)$$

where, $\left\langle (\Delta R_i)^2 \right\rangle$ is the mean squared fluctuations of residue i, and B_i is the measured quantity referred to as the B factor of residue i. The values of the B-factors are given in the Protein Data Bank.[1] The fluctuations of the junctions are correlated in space and time. The protein does perform its function through these correlated fluctuations.

To a first order approximation, the native protein may be considered as a non-random Gaussian network, with well defined mean position of each residue, and the fluctuations of two residues are under the effect of spring-like forces. The chiral carbon of each residue, called the alpha carbon or C^{α}, is taken as the reference atom of each residue, and the protein is considered in the coarse grained representation in terms of the C^{α}'s only. Two residues are said to be in contact if the distance between C^{α}'s is 7.0 Å, which is the radius of the first coordination shell of residues.

The statistical thermodynamics of random Gaussian networks is given by Flory.[14] The role of junction fluctuations on various properties of random amorphous networks was studied in detail, both theoretically and experimentally.[14–18] Based on these rubber elasticity models, a Gaussian Network Model (GNM) was formulated to study the fluctuations of residue positions in proteins.[19] This model was applied to the analysis of numerous proteins with significant success.[20–25] In Figure 1, we present a comparison of

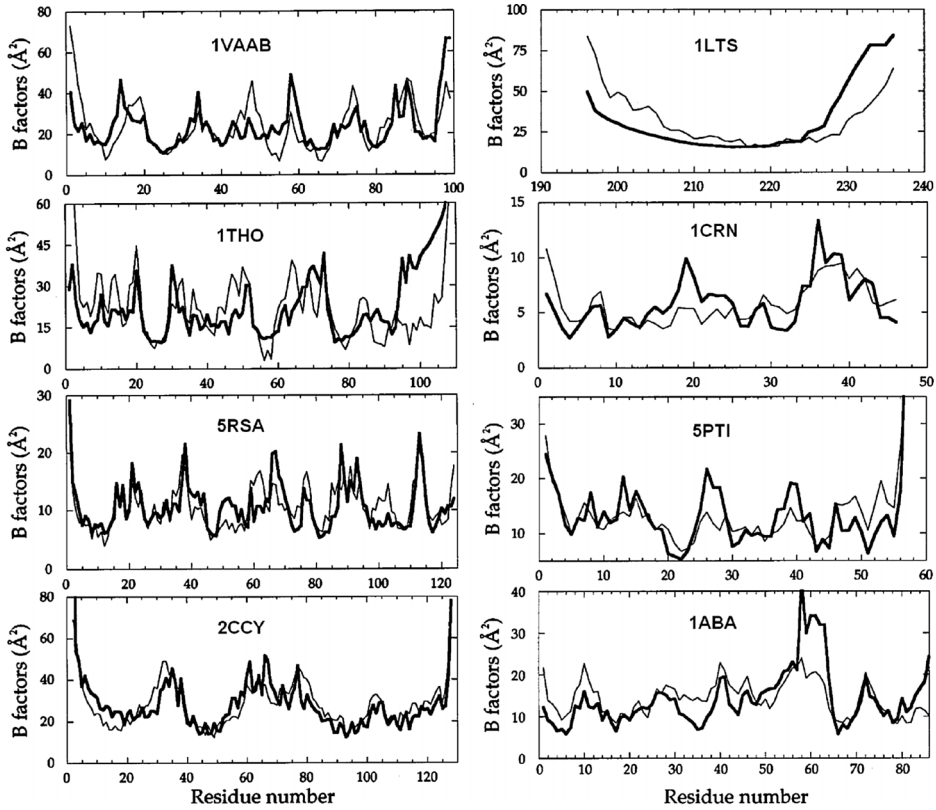

Figure 1.
Comparison of experimental (light lines) and theoretical (heavy lines) for various proteins. The Protein Data Bank identity of each protein is shown in each panel.

experimental data (light lines) with results of the theory based on the GNM.[19] Close agreement between the theory and experiment shows that the linear spring interactions between contacting residues explain the physics of the problem with reasonable accuracy.

Recently, the GNM was extended to the statistical thermodynamics analysis of proteins in aqueous media, and the correlations between energy fluctuations and residue fluctuations were characterized.[26,27] Coupling of the energy and residue fluctuations is of significant importance for determining the locations on a protein surface through which interactions with the external molecules are made.

In the sections below, we outline the model in some detail, and apply it to several proteins. First, we point out to the differences between random and nonrandom networks since function in a network is a direct consequence of nonrandomness.

Forming a Physical Network

Assume that there are m different species that will act as junctions. How can these come together to form a physical network? One can think of two ways of doing this:

(1) Form a linear, covalently bonded chain of a sequence of m beads, and fold the chain into its lowest free energy 3-d conformation. This is how proteins form in nature. The formation of the 3-d shape from a linear chain is called protein folding and is a rapid process of self organization.
(2) Allow the m unattached beads to collapse and self organize into a three dimensional structure.

Here we consider physical networks formed by the first scheme.

In systems that exhibit function, there are distinctive topological features and topology and function are closely related. We consider the topology in terms of the connectivity matrix, which is defined as follows: We number the repeat units, or the beads, from 1 to n. If unit i neighbors unit j in space, then we set the ijth entry of the connectivity matrix to unity. Otherwise it is zero. The Γ matrix of the Gaussian model of rubber elasticity,[14] which is the force constant matrix of the system of linear springs is then obtained from the connectivity matrix according to

$$\Gamma_{ij} = \begin{cases} -\gamma * & i \neq j \ \ and \ \ R_{ij} \leq r_{cutoff} \\ 0 \ \ i \neq j & and \ \ R_{ij} > r_{cutoff} \\ -\sum_{k} \gamma * & i = j \neq k \end{cases} \tag{2}$$

where, $\gamma *$ is the spring constant, taken equal for all interactions, and r_{cutoff} is the cutoff distance below which two junctions are assumed to be in contact.

The eigenvalue spectra of the Γ matrices for all proteins collapse approximately onto a single curve, shown in Figure 2 by the heavy solid curve. Now we assume physical networks that are formed either according to case (1) or (2). In case (1), the linear chain forms secondary structures such as helices and beta strands, and if it is then folded into a minimum energy 3-d structure, then the unique eigenvalue spectrum of the Γ matrix emerges. If a physical network is formed randomly, according to case (2) for example, then the eigenvalues spectrum is represented by the light curve in Figure 2. The shape of the light curve may be distorted by forming cross-links selectively, such as hydrophobic ones

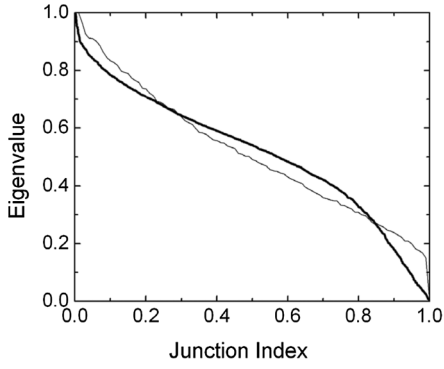

Figure 2.
Eigenvalues spectra of proteins and random networks. Heavy line is for native proteins, light curve for a random network.

favoring hydrophobic ones, but the spectrum of a protein, shown with the heavy curve, is never reached. The special form of the spectrum of a protein results from different levels of structural correlations resulting from the spatial packing requirements of helices, beta strands, turns, tight turns, loops, etc., and this level of correlation is necessary for function.

Interactions between Junctions and Correlations

We assume a system at equilibrium where the junctions exhibit large scale fluctuations. Since the three dimensional structure is intact, two junctions cannot wander away from each other. What holds the junctions together in a protein is the system of nonbonded interactions. What holds the junctions of an elastomeric chemical network together is the set of network chains.

In a protein, the number density of junctions vary from point to point. This variation is not random however, the nonrandomness resulting from the constraints imposed by the secondary structures within the protein domain. By default, a junction at the surface has fewer neighbors than one inside. However, the junction density at the surface changes strongly, mostly being very high at points where the protein forms a catalytic function. In Figure 3, we see the protein Carbon Monoxide Dehydrogenase, with the Protein Data Bank name of 1SU7.pdb. Figure 3a shows the all atom form, 3b shows the organization of the secondary structures, helices and beta strands.

The number of neighbors of each junction is shown in Figure 4 as a function of residue index where residues are sequentially numbered along the primary chain from one end to the other.

One sees that the average junction density of a protein as a physical network is comparable to an elastomeric network with about eight functional junctions.

The junctions of an elastomer fluctuate under the effect of network connectivity.[14] In a phantom network, the fluctuations of a given junction are related to the number of neighboring junctions. The crowded the neighborhood, the smaller the fluctuations. In fact the mean square fluctuations of junctions in a phantom network are related to the number of neighboring junctions as

$$\left\langle (\Delta R)^2 \right\rangle = \frac{(\phi - 1)}{\phi(\phi - 2)} \langle r^2 \rangle_0 \tag{3}$$

Here, $\left\langle (\Delta R)^2 \right\rangle$ is the mean square fluctuation of a junction, ϕ is the number of chains to which the given junction is attached, and, $\langle r^2 \rangle_0$ is the mean square

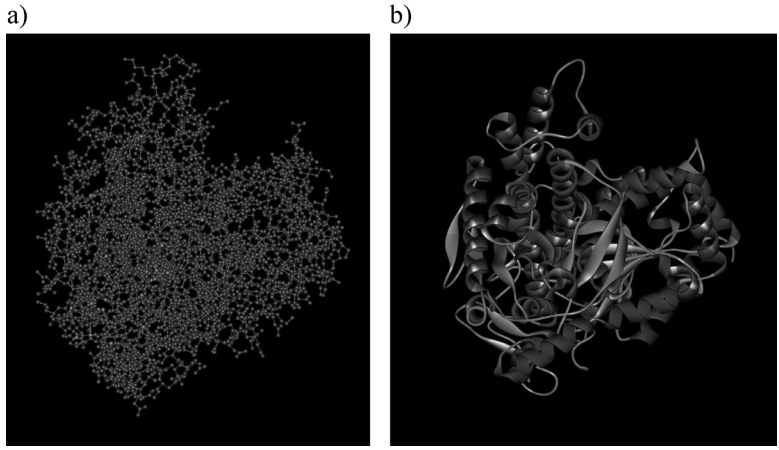

Figure 3.
(a) All atom depiction of the protein Carbon Monoxide Dehydrogenase, 1SU7.pdb. (b) The organization of secondary structure of the protein.

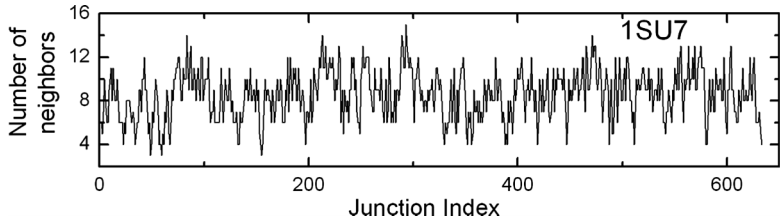

Figure 4.
Number of neighbors of junctions of the protein 1SU7.pdb.

dimensions of a chain between two junctions. The Gaussian network Model of proteins assumes that residues that are neighboring in space are connected by a Gaussian spring, and in the simplest theory, all spring constants are equivalent. Thus, this fictitious spring becomes the analogue of the Gaussian chain of rubber elasticity, and the residues become the junctions. In the remainder, junctions and residues will be used exchangebly.

Knowing the number of neighboring junctions should give an estimate of the fluctuations of junctions according to Eq. 3. In Figure 5, in the left panel, we compare predictions of Eq. 3 and experimental data for the protein 1SU7. There is some but not a very good agreement between theory and experiment. What is missing in the physics is that the correlations between the fluctuations of different junctions, such as $\langle \Delta R_i \cdot \Delta R_j \rangle$, where $i \neq j$ are not taken into account. The right panel shows the improvement when correlations are introduced (See the discussion of the following section for the right panel). To do this, we consider a statistical thermodynamic analysis of the system.

Statistical Thermodynamics of Fluctuations in Networks

We model a network, or a nano-gel embedded in a liquid. The system exhibits energy, volume and junction fluctuations. The probability distribution $f(\hat{U},\hat{V},\hat{\mathbf{R}})$ of the instantaneous values, $\hat{U}$, $\hat{V}$, and $\hat{\mathbf{R}}$, of the energy, volume and junction positions is given by the statistical - thermodynamic expression [28]

$$f\left(\hat{U},\hat{V},\hat{\mathbf{R}}\right) = \exp\left\{-k^{-1}\left[S - \frac{U}{T} - \frac{P}{T}V + \frac{\mathbf{F}}{T}\cdot\mathbf{R}\right] - k^{-1}\left(\frac{\hat{U}}{T} + \frac{P}{T}\hat{V} - \frac{\mathbf{F}}{T}\cdot\hat{\mathbf{R}}\right)\right\} \quad (4)$$

where, k is the Boltzmann constant, S is the entropy, P the pressure, T temperature and $\boldsymbol{F}$ is the force. The variables with a hat

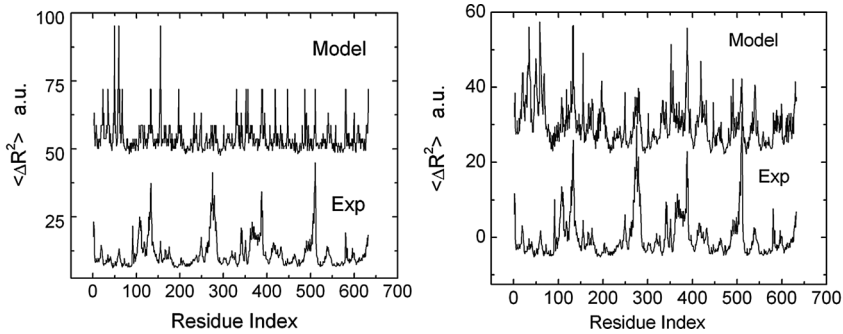

Figure 5.
Comparison of junction fluctuations. The lower curve is from experiments. The upper curves are obtained from Eq. 3 (Left panel), and inverse of the Γ matrix (Right panel).

are instantaneous variables, and those without are thermodynamic averages. The correlation of fluctuations $\langle \Delta \mathbf{R}_i \Delta \mathbf{R}_j{}^T \rangle$ of the ith and jth junctions are defined as

$$\langle \Delta \mathbf{R}_i \Delta \mathbf{R}_j{}^T \rangle = \sum \left(\hat{\mathbf{R}}_i - \mathbf{R}_i \right) \left(\hat{\mathbf{R}}_j - \mathbf{R}_j \right)^T \mathrm{f}\left(\hat{\mathrm{U}}, \hat{\mathrm{V}}, \hat{\mathbf{R}} \right) \quad (5)$$

Using Eq. 4 in Eq. 5 leads to the correlation

$$\langle \Delta \mathbf{R}_i \Delta \mathbf{R}_j{}^T \rangle = \mathrm{kT} \left(\frac{\partial \mathbf{R}_i}{\partial \mathbf{F}_j} \right)_{\mathrm{T,P,Fi} \neq \mathrm{j}} \quad (6)$$

In general, if Φ_k represents any of the extensive variables ΔU, ΔV, ΔR, and Ψ_k represent the conjugate variables $1/T$,- P, F, then, in principle, all higher moments of the extensive variables can be derived iteratively according to the rule[28]

$$\langle \phi \Delta \Phi_k \rangle = -k \frac{\partial}{\partial \Psi_k} \langle \phi \rangle + k \left\langle \frac{\partial \phi}{\partial \Psi_k} \right\rangle \quad (7)$$

where, ϕ denotes the fluctuations of the extensive variables, ΔU, ΔV, ΔR, or their product of any order. For example, letting $\phi = \Delta X_i$ and $\Delta \Phi_k = \Delta X_j{}^T$, since $\langle \Delta X_i \rangle = 0$, and

$$\left\langle \frac{\partial \Delta X_j}{\partial F_j} \right\rangle = \left\langle \frac{\partial \hat{X}_j}{\partial F_j} \right\rangle, \quad (8)$$

we obtain

$$\langle \Delta X_i \Delta X_j{}^T \rangle = kT \left(\frac{\partial \widehat{X}_i}{\partial F_j} \right)_{\mathrm{T,P,Fi} \neq \mathrm{j}} \quad (9)$$

Higher order moments can also be obtained by a recursion relation.[28]

If we assume that the fluctuations result from a harmonic interaction between the junctions, then the force-position equation of state is $F = \Gamma \hat{R}$ where, Γ_{ij} is the spring constant matrix, defined by Eq. 2. Substituting $F = \Gamma \hat{R}$ into Eq. 9 leads to

$$\langle \Delta R_i \Delta R_j \rangle = kT \left(\Gamma^{-1} \right)_{ij} \quad (10)$$

We can similarly show the correlation of the energy fluctuations with fluctuations of residue positions. This is important because a strong correlation $\langle \Delta U \Delta \mathbf{R}_i \Delta \mathbf{R}_j{}^T \rangle$ between the total energy uptake of the protein and the residues i and j directly points to the residues i and j that are active in this energy transfer. For this, we write the energy fluctuations for the harmonic system as

$$\Delta U = F_k \Delta R_k = F_k F_j \left(\Gamma^{-1} \right)_{jk} \quad (11)$$

Here, summation is understood over repeated indices. This expression is based on the assumption that the system is in a state of ease when the residues are at their mean positions.

Choosing $\phi = \Delta U \Delta R_i$ and using the recursion relation, Eq. 7, we have

$$\langle \Delta U \Delta R_i \Delta R_j \rangle = kT \frac{\partial}{\partial F_j} \langle \Delta U \Delta R_i \rangle + kT \left\langle \frac{\partial}{\partial F_j} (\Delta U \Delta R_i) \right\rangle \quad (12)$$

Rearranging the terms, as outlined in Reference[29] leads to the energy-fluctuation correlation

$$\langle \Delta U \Delta \mathbf{R}_i \Delta \mathbf{R}_j{}^T \rangle = kT \langle \Delta \mathbf{R}_i \Delta \mathbf{R}_j{}^T \rangle = (kT)^2 \left(\Gamma^{-1} \right)_{ij} \quad (13)$$

Equations 10 and 13 are the main result of the statistical thermodynamic model. The first is known as the Gaussian Network Model (GNM) of proteins, and has been successful in predicting correlations and function in proteins. The diagonal elements $\langle \Delta R_i{}^2 \rangle = kT \left(\Gamma^{-1} \right)_{ii} = \frac{3}{8\pi^2} B_i$ are proportional to the B factors and can be compared with experimental data. In the right panel of Figure 5, we show the results for 1SU7, where the upper curve is obtained from the inverse of the Γ matrix. We now see that the agreement with experiment and theory is much improved. Several other examples showing good agreement with experiment are also given in Figure 1.

Figure 5 shows that number of neighbors is not a sufficient criterion for explaining the fluctuations of junctions. The Gaussian Network Model not only predicts the fluctuations of residues, but also motions of different parts of the protein relative to each other. For this purpose, we write the

correlation matrix in modal form and look at the slow modes that characterize the large scale motions.

The correlation matrix may be expressed in modal form as[30]

$$\langle \Delta R_i \Delta R_j^T \rangle = \sum_k \lambda_k^{-1} \left[e_k e_k^T \right]_{ij} \quad (14)$$

where, λ_k is the kth eigenvalue of the Γ matrix, e_k is the corresponding eigenvector, and $[]_{ij}$ is the ijth element of the enclosed matrix. Keeping the small eigenvalues in the sum on the right hand side of Eq. 14 gives information on the global features of the protein structure. For example, parts of the protein that moves as individual domains are characterized by the small eigenvalues. The example of HIV Reverse Transcriptase, which is a complex that has a flexible and a rigid part, has been studied in terms of the smallest few eigenvalues and eigenvectors of the Γ matrix.[21] With this example, and several others following this work, function that is associated with global motions of a network is shown to derive from the slow modes of its topology matrix.

Large eigenvalues of the Γ matrix are also of physical significance and strongly related to function. The largest eigenvector of the connectivity matrix points out to nodes that can be perturbed the most. We regard the term 'perturbation' as the process of transferring energy to the system from an outside agency. Thus, we consider Eq. 12, which tells us about the residues in a protein that can exchange energy with the surroundings and with other residues of the protein. This equation is very important because it identifies the points of a system that can interact with the environment, and transfer the input from the environment to other points of the system. This is essentially how the system performs its function.

Summing both sides of Eq. 12 over the jth index leads to the total coupling $C_{T,i}$ of residue i to its surroundings

$$C_{T,i} = \sum_j \langle \Delta R_i \Delta R_j^T \rangle = kT^{-1} \sum_j \langle \Delta U \Delta R_i \Delta R_j^T \rangle \quad (15)$$

The last term in Eq. 15 acknowledges the role of energy exchange of residue i with its surroundings that consist of the neighboring residues and the surroundings of the protein. Thus, $C_{T,i}$ is a measure of the energy exchange of the ith node with its surroundings, where the term surrounding now includes the environment and the other junctions of the system.

An example using Eqs. 12 and 15 is the Heme oxygenase (HO) which is responsible for the degradation of heme to biliverdin. In the heme bound state, Human heme-oxygenase-1 (HO-1) arranges its helical shape with the help of highly conserved helix residues, so that it supplies flexibility to accommodate substrate binding and product release.[31] Human HO-1 has a dynamic active-site pocket, which is enlarged in the empty state as the helices surrounding the heme plane move farther apart. In the presence of the heme, the residues Thr21, Val24, Thr23, Thr26, Ala28, Glu29, Tyr-134, Thr-135, Leu-138, Gly-139, Ser-142, and Gly-143 interact with heme.[32,33]

Figure 6 shows the total correlation, C_T, of a given residue, presented as the residue index along the abscissa, obtained by using 1NI6.pdb. In Figure 6b, the residues that exchange energy with the surroundings are identified with a darker hue. The heavy vertical strip shows that the residues 118-124 interact with all the residues of the protein.

Manipulating the Function of Proteins: Drug Design

A major technological effort for controlling the function of a protein is drug design. Drug design has the following general steps:

1. Identify the gene that causes the disease, and the protein that is affected by this in the form of mutation.
2. Identify the nodes on the protein that can exchange energy with this mutated gene (Druggable regions).
3. Decide on one or more of these nodes to which a drug will be attached to affect the function of the mutated protein.

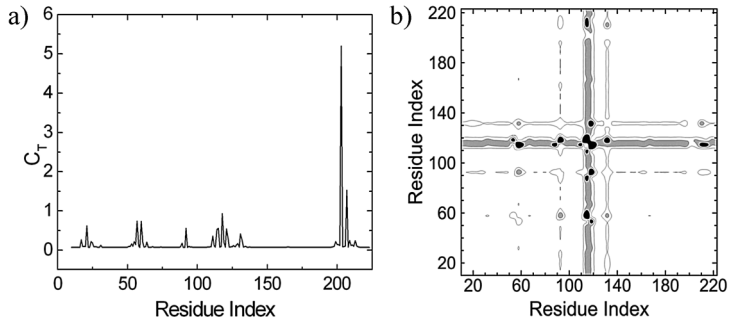

Figure 6.
(a) Total correlation (b) and the contour plot for fluctuations of Heme oxygenase.

4. Design a small molecule that will attach to this node by establishing favorable interactions with these points.

The program stated above corresponds to the well known recognition problem of polymer physics. Given a patch on the surface of a globular collapsed polymer, find a molecule that will recognize this patch. The simple model outlined above, in terms of the fast modes of the GNM have been shown by us to predict the drug binding sites.[26,27] A functional nano-gel can be manipulated in a similar way, on which our group is working.

Conclusion

As reviewed briefly in the Introduction section, significant effort goes into building functional synthetic gels. Here, we showed that a simple harmonic model can predict function in proteins regarded as physical gels, based purely on topological properties. If the physical gel can be designed with a planned topology, it should be possible, in principle, to obtain a structure that will perform the desired nano-scale fluctuations. As implied in the introduction section, a necessary criterion for this is the presence of secondary structure that controls the inhomogeneous local structure. Building secondary structure into synthetic polymers may not be a straightforward task however, unlike the case of peptides where strong propensities for helical, extended and loop structures exist in the polypeptide chain. The chemical nature of functional groups is important in the systems discussed, yet we based the model only on topological considerations that do not relate to the chemical identity at first sight. However, the topology of the protein is essentially an evolutionary result of the chemical structure built into the primary chain which in turn determines the three dimensional structure.

The emphasis in this paper has been on physical networks, mimicking native proteins. There is an emerging alternative technique for control of function by incorporating chemically adaptable junctions into the system.[34] Although physically adaptable junctions are widely chosen by biological systems, control of topology by manipulating the junctions covalently is promising, particularly in redox manipulation.

One important pathway for function is allosteric control where an external agent, such as a drug molecule or a change in pH, operates on one part of the molecule that induces an action on another part. The Gaussian Network Model presented here has been successful in predicting allosteric sites.[26,27] Changes in pH induce local changes in the contact order of a few junctions, and this change in turn correlates with the fluctuations of other junctions at another point in the molecule. Here, we showed that such correlations may be predicted by analyzing the correlation

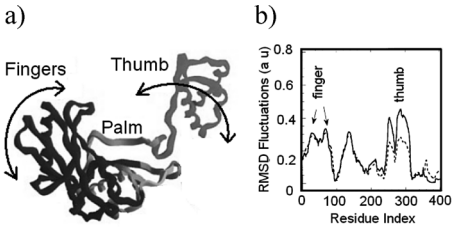

Figure 7.
(a) Collective motions, shown by the curved arrows, of the HIV-1 Reverse Transcriptase.[21] (b) Experimentally observed (dotted curve) RMSD fluctuations predicted by the slowest two modes (solid curve) of the Γ matrix.

matrix corresponding to the largest eigenvalues of the topology matrix. The resulting large scale motions such as motion of various domains of the network can in turn be predicted by analyzing the correlation matrix corresponding to the smallest eigenvalues. An example is the analysis of the collective motions of the protein HIV-1 Reverse Transcriptase[21] where the opening-closing motions of the protein regions can be predicted by the smallest nonzero eigenvalues as shown in Figure 7. In the left panel of Figure 7, the motions of a 'hand-like' region is depicted. On the right panel, the rmsd fluctuations for the fingers and the thumb obtained by the smallest to nonzero eigenvalues are shown.

Acknowledgements: This work was partially supported by the Turkish Academy of Sciences.

[1] Protein Data Bank http://www.rcsb.org/pdb/home/home.do.
[2] Y. Osada, A. R. Khokhlov, Eds., *Polymer Gels and Networks*, **2002**, Marcel Dekker, New York
[3] N. M. Sangeetha, and U. Maitra, *Chemical Society Reviews*, **2005**, *34*(10), 821–836.
[4] M. Shibayama, T. Norisuye, S. Nomura, *Macromolecules*, **1996**, *29*(27), 8746–8750.
[5] B. Xu, *Langmuir*, **2009**, *25*(15), 8375–8377.
[6] H. Maeda, *Chemistry-a European Journal*, **2008**, *14*(36), 11274–11282.
[7] P. Borckmans, et al., ed. *Chemomechanical Instabilities in Responsive Materials*, **2007**, Springer.
[8] F. Carpi, E. Smela, Eds., *Biomedical Applications of electroactive Polymer Actuators*, **2009**, Wiley.
[9] S. K. Ahn, et al., *Soft Matter*, **2008**, *4*(6), 1151–1157.
[10] M. Irie, *Microchemistry*, **1994**, *571*, 363–371.
[11] Y. Osada, and J. P. Gong, *Progress in Polymer Science*, **1993**, *18*(2), 187–226.
[12] F. Liu, and M. W. Urban, *Progress in Polymer Science*, **2010**, *35*(1–2), 3–23.
[13] C. Branden, and J. Tooze, *Introduction to protein structure*, **1999**, New York Garland.
[14] P. J. Flory, *Proc. R. Soc. London, A*, **1976**, *351*, 351.
[15] B. Erman, and P. J. Flory, *Macromolecules* **1982**, *15*, 806–812.
[16] B. Erman, and P. J. Flory, *Macromolecules*, **1985**, *19*, 2342.
[17] P. J. Flory, and B. Erman, *Macromolecules*, **1982**, *15*, 800–806.
[18] B. Erman, J. E. Mark, *Structures and Properties of Rubberlike Networks*. **1997**, New York Oxford University Press.
[19] I. Bahar, A. R. Atilgan, and B. Erman, *Folding & Design* **1997**, *2*, 173–181.
[20] I. Bahar, et al., *Phys. Rev. Lett.*, **1998**, *80*, 2733–2736.
[21] I. Bahar, et al., *J. Mol. Biol.*, **1999**, *285*, 1023–1037.
[22] A. Erkip, and B. Erman, *Polymer*, **2004**, 641–648.
[23] B. Erman, *Biophysical Journal*, **2006**, *91*, 3589–3599.
[24] T. Haliloglu, I. Bahar, and B. Erman, *Phys. Rev. Lett.*, **1997**, *79*, 3090–3093.
[25] O. N. Yogurtcu, M. Gur, and B. Erman, *Journal of Chemical Physics*, **2009**, *130*(9), e095103.
[26] T. Haliloglu, and B. Erman, *Physical Review Letters*, **2009**, *102*, 088103–088106.
[27] T. Haliloglu, E. Seyrek, and B. Erman, *Physical Review Letters*, **2008**, *100*, 228102–4.
[28] H. B. Callen, *Thermodynamics and an introduction to thermostatistics*, Second ed. **1985**, Wiley.
[29] T. Haliloglu, A. Gul, and B. Erman, *Plos Computational Biology*, **2010**, *6*(7), e1000845.
[30] I. Bahar, et al., *Phys. Rev. Lett.*, **1998**, *80*, 2733–2736.
[31] D. J. Schuller, et al., *Nature Structural Biology*, **1999**, *6*(9), 860–867.
[32] L. Lad, et al., *J Biol Chem*, **2003**, *278*(10), 7834–43.
[33] G. N. LaMar, et al., *J Biol Chem*, **2001**, *276*(19), 15676–87.
[34] C. J. Kloxin, et al., *Macromolecules*, **2010**, *43*(6), 2643–2653.

