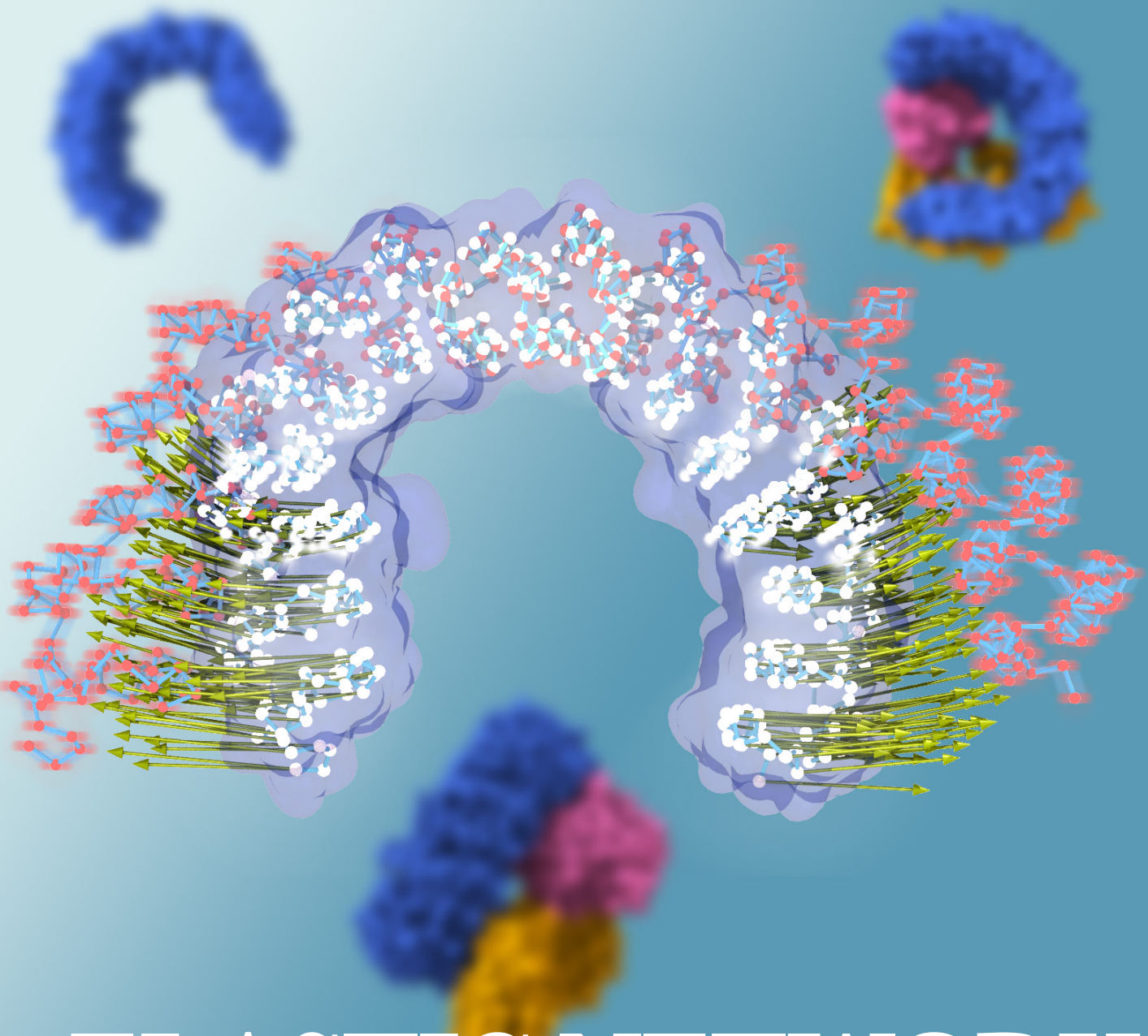

ELASTIC NETWORK MODELS IN PROTEIN DYNAMICS:

Bridging Simplicity and Complexity

By: Mert Gur, Mert Golcuk, Ivet Bahar

https://doi.org/10.51167/acm00072

Elastic Network Models (ENMs) have improved our understanding of protein dynamics by providing a simplified yet accurate representation of complex molecular motions. Originating from the fundamental theory and methods of solid-state physics and polymer science, ENMs approximate biomolecular structures, both proteins and DNA/RNA molecules and their complexes, as networks of nodes (residues) connected by elastic springs. This simple representation allows for obtaining an analytical solution for the spectrum of motions uniquely accessible to the structure, similar to that obtained by normal mode analysis (NMA) of molecular systems. This review traces the evolution of ENMs from their origin, highlighting key developments such as the Gaussian Network Model (GNM) and the Anisotropic Network Model (ANM) originally introduced in the late 1990s and early 2000s by Turkish scientists at Bogazici University Polymer Research Center (PRC). These models have substantially impacted the field by offering insights into how protein flexibility, collective changes in conformation, and coupling within and across monomers/subunits in large assemblies define the mechanisms of motions that often enable the biological actions. We provide a brief overview of the theoretical foundations of ENMs, and the simplicity and computational efficiency of the underlying methods, along with illustrations of its use in predicting residue fluctuation profiles or global changes in structures. Additionally, we emphasize the integration of ENMs with all-atom molecular dynamics (MD) simulations and cryo-electron microscopy (cryo-EM) data, showcasing their utility of these hybrid modeling approaches. The review also discusses the utility of ENMs in elucidating the effects of oligomerization and membrane interactions on the functional dynamics of proteins. Through these discussions, we underscore the critical role and expanding potential of ENMs in computational biophysics and structural biology.

Mert Gur
Assoc. Prof. Mert Gur is the Executive Director of the Computational Biomedicine & Biotechnology M.S. Program at the University of Pittsburgh (Pitt) School of Medicine (SOM) Department of Computational and Systems Biology (CSB) and a faculty at Istanbul Technical University (ITU) Mechanical Engineering (ME) Department. He earned his B.S. in ME (2006) at the Middle East Technical University and his Ph.D. in Computational Science and Engineering (2010) at Koc University in Türkiye. He worked as a postdoctoral associate at the Pitt SOM CSB and then at Lawrence Berkeley National Laboratory, until joining ITU ME as a faculty member in 2015. He served as Vice Dean of the School of ME at ITU (2018-2020) and as Vice Dean of the Graduate School of Science, Engineering, and Technology (2020-2022) at ITU. His research focuses on Computational Structural Biology, Computational Biomedicine, and Mechanical Engineering. He received several prestigious awards in Türkiye.

Mert Golcuk
Mert Golcuk holds a B.Sc. and M.Sc. from Istanbul Technical University, where he is currently pursuing his Ph.D. under the supervision of Assoc. Prof. Mert Gur. His research leverages molecular dynamics simulations to explore the mechanisms of molecular motors, membrane-associated proteins, the SARS-CoV-2 spike protein, and nanobody inhibition. His work also delves into drug targeting and peptide design.

Ivet Bahar
is currently Louis & Beatrice Laufer Endowed Professor and Director at the Laufer Center for Physical and Quantitative Biology of Stony Brook University. She was a professor at Bogazici University Chemical Engineering Department (1986-2001) and the University of Pittsburgh (2001-2022) before that. Her major area of research is theory and computations of biochemical and biophysical systems, with applications to protein dynamics and drug discovery. She is an elected member of European Molecular Biology Organization (EMBO) and the US National Academy of Sciences (NAS).

History of Elastic Network Models: From Complex Energy Surfaces to Elegant Simplicity

Proteins possess a complex energy landscape characterized by a rugged, funnel-like shape with multiple minima.[1,2] Motions within a minimum are mainly harmonic,[3-5] while the anharmonic components arise from transitions between minima.[3] Thus, the energy landscape can be approximated by a harmonic shape near their global energy minima (Figure 1). This foundational observation has facilitated the application of normal mode analysis (NMA),[6] which simplifies the computational modeling of protein dynamics by concentrating on the harmonic portion of the energy surface during equilibrium motions. Introduced at the atomic level in the 1980s through the pioneering work of Go et al.[7], Brooks and Karplus[8], and Levitt et al.[9], NMA offers a computationally efficient approach to exploring the space of accessible motions near equilibrium conditions; whereas all-atom molecular dynamics (MD) simulations may provide additional insights by accounting for anharmonic contributions and exploring multiple energy minima. A harmonic potential well in NMA has the following form,[6]

$$U(\mathbf{r}) = \frac{1}{2}(\mathbf{r}-\mathbf{R})^T \mathbf{H}(\mathbf{R})\,(\mathbf{r}-\mathbf{R}) \qquad (1)$$

$\mathbf{R}$ is a 3N-dimensional vector describing the stable conformation at the center of the well, where N is the number of atoms. Similarly, $\mathbf{r}$ is a 3N-dimensional vector that represents the current/instantaneous protein conformation. The superscript *T* denotes the transpose, i.e. the row vector ($\mathbf{r}$-$\mathbf{R}$) in this case. The matrix $\mathbf{H}$, known as the Hessian, is symmetric and positive semidefinite matrix, and characterizes the shape of the potential well. Given a potential energy function *U*, each element of $\mathbf{H}$ is calculated as the second derivative of *U* with respect to pairs of, coordinates q_i, q_j ($1 \le i,j \le 3N$; e.g. the *x-, y-, and z-*components of the atomic position vectors), evaluated at the local minimum $\boldsymbol{R}_{min}$.

$$\mathbf{H}_{ij} = \left[\frac{\partial^2 U}{\partial q_i \partial q_j}\right]_{\mathbf{r}=\mathbf{R}_{min}} \qquad (2)$$

Yet, the requirement for a detailed all-atom force field, combined with the complexity and mathematical challenges associated with energy minimization to locate $\mathbf{R}_{min}$, has limited the broader adoption of NMA. Furthermore, while the harmonic approximation may provide a good description of the energy landscape near an energy minimum, the complete landscape may have multiple minima, especially when modeled at full atomic scale (as illustrated in Figure 1 for dynein, a motor protein[10]), which limits the applicability of NMA at full atomic scale.

Starting from 1996, a classical paper published by Monique Tirion in *Physical Review Letters* led to a new era: Tirion[11] demonstrated that a full atomic NMA using harmonic potentials with uniform spring constants for all atom pairs within an interaction range could yield low frequency mode shapes nearly identical to those deduced from NMA carried out with full-fledge force fields typically used in MD simulations. The insensitivity of the low frequency modes to the force field, and the possibility of adopting uniform harmonic potentials (or elastic springs) between all atom pairs to end up with mode shapes almost indistinguishable from those otherwise obtained with sophisticated potentials and tedious energy minimization algorithms, caught the attention of the Turkish scientists at the PRC (Bogazici University, Istanbul). Rather than carrying an NMA, the scientists Bahar, Erman, and Atilgan came up with an analytical model[12] inspired by the classical Rouse model of polymer physics. According to the Rouse model, the macromolecule is represented by a string of beads connected by springs of uniform force constant. The beads represent sequential structural units (monomers or segments of polymeric chain), such that the springs connect sequential neighbors (only) along the chain. The overall connectivity of the macromolecule is then represented by the so-called Rouse matrix of the form

$$\Gamma_{ii}(Rouse) = \begin{cases} 1 & if\ \ i = 1\ and\ i = N \\ 2 & if\ \ 2 \le i \le N-1 \end{cases}$$

$$\Gamma_{ij}(Rouse) = \begin{cases} -1 & if\ |i-j| = 1 \\ 0 & otherwise \end{cases} \quad (3)$$

The PRC team, inspired by Tirion's paper and fundamental concepts of solid-state physics, as well as the theory of elasticity of polymer networks[13,14], came up with the idea of replacing the Rouse matrix (of polymers) by a new form, the Kirchhoff matrix Γ, with the major difference of connecting not only the sequential neighbors along the chain, but also those making spatial contacts. The structure would thus be represented by a 'network', rather than a spring-and-beads model. This way Γ would account not only

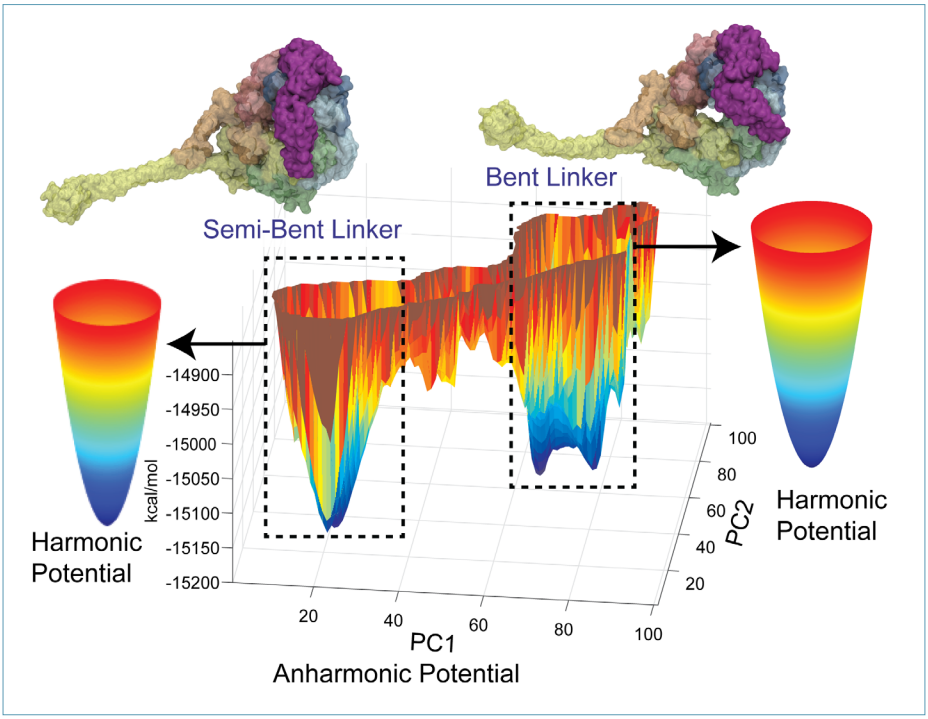

Figure 1: Anharmonic potential energy surface associated with the priming stroke of dynein and approximately harmonic potentials near the end points. *Middle panel:* **The potential energy landscape of dynein motor protein priming stroke obtained from 11,250 ns of all-atom MD simulations[10]. Two minima are sampled corresponding to the dynein's straight and semi-bent conformations.** *Left and right panel*: **Approximating the energy surface near the semi-bent and straight linker states by a harmonic energy surface. Dynein is responsible for intracellular transport, and its linker domain plays a crucial role in generating the "priming stroke" a conformational change that prepares the motor for its power stroke and provides forward swinging motion, facilitating movement along microtubules.**

for chain connectivity but also tertiary contacts. Tertiary contacts indeed are the major features that distinguish biological polymers (proteins and DNA) from synthetic polymers. The Γ matrix then becomes

$$\Gamma_{ii} = \{z_i \quad if\ 2 \le i \le N-1$$

$$\Gamma_{ij} = \begin{cases} -1 & if\ r_{ij} < r_{cut} \\ 0 & otherwise \end{cases} \quad (4)$$

where z_i is the coordination number of node i, given by the negative sum of all non-diagonal elements in the same row (or column, since Γ is symmetric), r_{ij} is the distance between nodes i and j, and r_{cut} is a cutoff distance for defining node pairs that are connected by a spring. In this new model, called the Gaussian Network Model (GNM) introduced by Bahar, Atilgan, and Erman in 1997,[12] the nodes represent the individual amino acids, their locations being identified by that resolved by X-ray crystallography (or NMR or cryo-EM) and the springs account for those connected or making secondary and tertiary contacts, by adopting a cutoff distance of r_{cut}=7.0 Å. This distance is representative of a first shell of inter-residue coordination based on the extensive evaluation of the contact probabilities of residue pairs[15] in structures deposited in the Protein Data Bank (PDB). The overall potential of the protein represented by the GNM is given by

$$U(\Delta\mathbf{R}) = \frac{\gamma}{2}(\Delta\mathbf{R})^T\mathbf{\Gamma}(\Delta\mathbf{R}) \quad (5)$$

where γ is a uniform force constant for all springs of the model. Equation (5) has same form as Equation (1); $\Delta\mathbf{R}$ represents in this case the change in the position of the N residues, $(\Delta\mathbf{R})^T = (\Delta R_1\ \ \Delta R_2\ \ \Delta R_3\ \ \ldots\ \ \Delta R_N)$ away from their equilibrium positions[12,16]. GNM allows for a simple evaluation of normal modes in the N-dimensional space of residues[16]. Elastic springs are defined only between C_α atoms that are within a cutoff distance of r_{cut}. This connectivity information is stored in the Kirchhoff matrix $\mathbf{\Gamma}$ by setting diagonal elements to the degree of each node and off-diagonal elements to -1 if connected, as described by Equation (4). The fluctuations are assumed to be isotropic and Gaussian, such that the model provides a description of the size of motions (e.g. the mean-square fluctuations $<(\Delta\mathbf{R})_i^2>$ of amino acids, or their cross-correlations $<\Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j>$), not their directionality. The mean-square (ms) fluctuations of the individual residues and their cross-correlations

are obtained by generalized Gaussian integrations over all fluctuations as

$$< (\Delta\mathbf{R})_i^2 > = \int (\Delta\mathbf{R})_i^2 \exp\{-U(\Delta\boldsymbol{R})/kT\}\, d\Delta\mathbf{R} / \int \exp\{-(\Delta\boldsymbol{R})/kT\}\, d\Delta\mathbf{R} = \frac{3kT}{\gamma}[\Gamma^{-1}]_{ii} \quad (6)$$

and

$$< \Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j > = \frac{3kT}{\gamma}[\Gamma^{-1}]_{ij} \quad (7)$$

Here k is the Boltzmann constant, T is the absolute temperature, and $[\mathbf{\Gamma}^{-1}]_{ij}$ denotes the ij^{th} element of the pseudoinverse of $\mathbf{\Gamma}$. The GNM lends itself to a unique evaluation of normal modes for each structure, upon eigenvalue decomposition of $\mathbf{\Gamma}$, as $\mathbf{\Gamma}=\mathbf{U}\Lambda\mathbf{U}^T$, where $\mathbf{U}$ is the N x N matrix of eigenvectors $[\mathbf{u}_0\ \mathbf{u}_1\ \mathbf{u}_2\ \ldots \mathbf{u}_{N-1}]$, and Λ is the diagonal matrix of eigenvalues comprised of the zero eigenvalue ($\lambda_0=0$), and non-zero eigenvalues organized in ascending order $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_{N-1}$, scaling with the frequency of individual modes. The N elements of the eigenvector $\mathbf{u}_k$ describe the normalized displacements of the N residues along the k^{th} mode. Often times, a few modes at the low frequency end of the spectrum (e.g., k=1-3), also called global modes, make a dominant contribution to the overall dynamics. The fluctuations profile of residues in these modes reflect the residue displacements in the most cooperative motions that embody the entire structure. The contribution of three slowest modes to $<(\Delta\mathbf{R})_i^2>$, for example, is given by

$$< (\Delta\mathbf{R})_i^2 >|_{k=1-3} = \frac{3kT}{\gamma}\sum_{k=1}^{3}(\frac{1}{\lambda_k})[\boldsymbol{u}_k\boldsymbol{u}_k^T]_{ii} \quad (8)$$

Previous applications have shown that such motions driven by the slowest mode often relate to cooperative displacements that are evolutionarily optimized/selected to achieve function; whereas those in the other end of the spectrum correspond to high frequency fluctuations undergone by the most severely confined (often high conserved, core) residues of the protein[17-20]. Therefore, to predict the modes of motion of a protein via GNM, only two simple steps are required: First, the connectivity information from the known structure in the PDB is transferred into matrix form, constructing the Kirchhoff matrix. Then, eigenvalue decomposition is performed directly on this matrix to separate the fluctuations into distinct modes of motion and evaluate the dynamic features driven by these modes, as in Equation (8).

GNM was the first in a series of elastic network models (ENMs), with others being proposed in the following 4-5 years and several variations developed since then. Of particular interest is the broadly used Anisotropic Network Model (ANM) introduced by Bahar and her collaborators[21]. Unlike GNM, which assumes an isotropic motion, ANM accounts for anisotropic motions within a 3N-dimensional configurational space. The Kirchhoff matrix is replaced in this case by a Hessian that can be readily expressed by a simply analytical expression (without the need to carry out energy minimization) using PDB coordinates to define equilibrium conformation[21]. Note that the original derivation of the ANM, based on a force balance at each node, differs from that of classical NMA, but the resulting expressions for the spectrum of normal modes are the same. The major difference is the use of uniform harmonic potentials for at the residue level in the ANM, as opposed to a full force field at atomic level in the classical NMA. The displacements of residues are therefore expressed in terms of their three components $\Delta\boldsymbol{R}_i=[\Delta x_i, \Delta y_i, \Delta z_i]$ leading to the ***3N***-dimensional eigenvectors and ***3N-6*** normal modes (excluding the rigid-body translation and rotations) described by the non-zero eigenmodes of the ANM Hessian. Therefore, the ANM provides information on the directions of motions or the so-called 3N-dimensional deformation vectors $\boldsymbol{d}^k$ of the overall structure driven by the individual modes k.

In the next sections, we illustrate the utility of ENMs in predicting structural dynamics in accord with complex experiments and simulations, how ENMs help us understand the molecular basis of functional interactions, such as the effects of oligomerization and coupling to membrane dynamics. Additionally, recent years have seen an explosion in the number of hybrid methodologies that combine ENMs and MD simulations, especially in the applications to modeling cryo-EM structures' dynamics, which we briefly summarize below.

Predicting Residue Fluctuations and Conformational Transitions with the ENMs

Temperature factors, also known as B-factors or Debye-Waller factors, describe the atomic displacement or thermal motion in a crystallographic structure, providing insights into the relative flexibility of protein residues. Even though the B-factors are biased by crystallographic contacts, and the observed motions are suppressed due to crystal environment (as compared to those observed by NMR in solution), the B-factors provides a first estimate of the relative flexibilities of individual residues. Likewise, the multiple models deposited for structures resolved by NMR spectroscopy reflect the breadth of conformational space accessible to proteins. Cryo-electron microscopy (cryo-EM) further offers insights into protein dynamics by revealing various classes of conformations. Many studies[12,21-25] have confirmed the effectiveness of various ENMs in predicting residue ms fluctuations in reasonable agreement with the B-factors despite the approximations in both theory and experiments.

Furthermore, while ENMs such as GNM and ANM, successfully predict experimentally obtained native state residue fluctuations,[12,21,24,25] they also satisfactorily reproduce residue fluctuations obtained with significantly more expensive all-atom MD simulations[26,27] (Figure 2). In addition to predicting residue fluctuations, ENMs have also proven to predict structural transitions in accordance with those observed in experiments[28,29] and sufficiently long MD simulations[4]. For example, the modes of motions predicted by the ANM were shown to agree with global motions exhibited by archaeal aspartate transporter, GltPh, in microseconds-long all-atom MD simulations, which were crucial for understanding the mechanism of transport of the neurotransmitter glutamate in the human central nervous system. Likewise, milliseconds-long MD simulations for bovine pancreatic trypsin inhibitor (BPTI), a model protein important for studying protein folding and stability relevant to various diseases, yielded fluctuation behavior nearly identical to that obtained by the ANM, while ANM results are obtained within several orders of magnitude shorter time. Furthermore, ANM-predicted collective motions along different modes were shown to agree with the conformational transitions undergone between the BPTI sub-states sampled in expensive MD simulations.[4]

Revealing the Influence of Oligomerization and Membrane Coupling on Functional Dynamics

ANM was utilized to elucidate the impact and underlying physical mechanisms of protein oligomerization, specifically focusing on neurotransmitter symporters (NSS family). These studies demonstrated that dimerization significantly enhances the transitions between the outward-facing open (OFo) and inward-facing open (IFo) states of leucine transporter (LeuT) and dopamine transporter (DAT), which are crucial for enabling substrate transport from the extracellular or synaptic region to the cytoplasm (Figure 3).[29] OFo and IFo refer to the conformational states that allow for alternating access to the

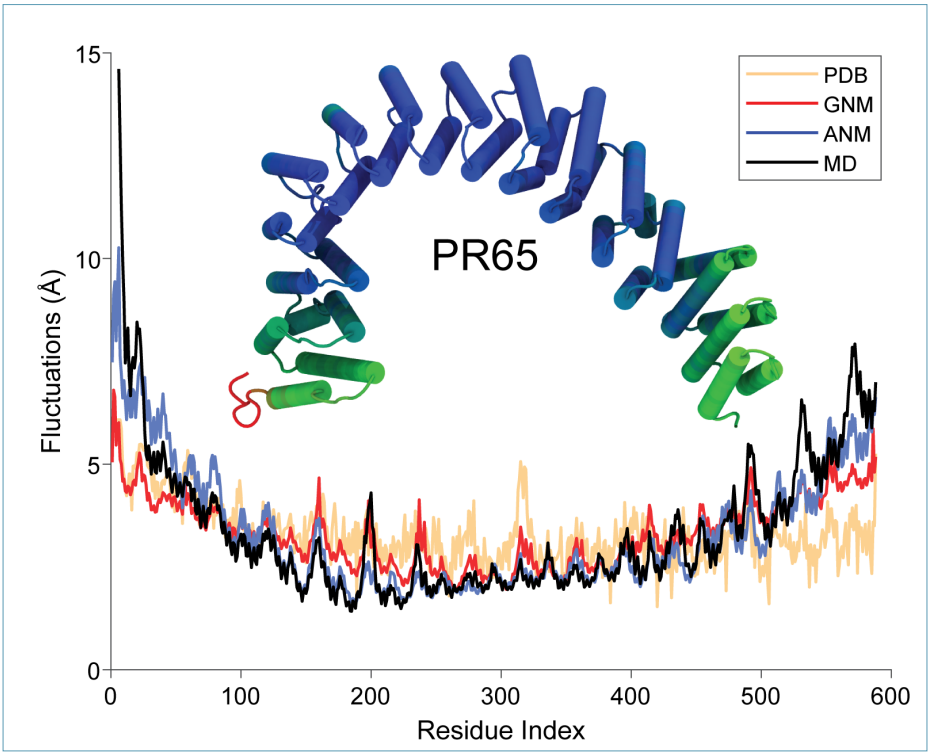

Figure 2: Residue fluctuations of PR65, the scaffolding subunit of the phosphatase PP2A. Residue fluctuation were obtained from GNM[30], all-atom MD simulations[31] and ANM analysis of the monomeric PR65, while the B-factors for PR65 are reported for the homo-dimeric PR65 X-ray structure (PDB ID 1B3U[32]). Magnitudes were normalized with respect to MD simulations. The PR65 structure (PDB ID: 6NTS[33]) is shown in a cartoon representation, with the magnitude of residue fluctuations observed in MD simulations color-coded from blue *(lowest)* to red *(highest,* at the N-terminus, with green indicating relatively large fluctuations. The periodicity of fluctuations originates from the tandem repeat (TR) structure of PR65 which is composed of 15 HEAT repeats.

extracellular (EC) and intracellular (IC) vestibules, respectively. These transitions enable the binding of substrates in the OFo state and their release in the IFo state, thereby completing the transport cycle. By examining the collective dynamics of monomers and dimers, the study revealed that dimerization induces specific modes of motion that facilitate these functional transitions. This marked the first instance of using ANM to provide a detailed mechanistic understanding of how oligomerization modulates the intrinsic motions of protein monomers, thereby playing a vital role in their biological function. The findings underscore the importance of protein oligomerization in achieving efficient substrate transport, providing valuable insights into the structural dynamics of NSS proteins and their functional optimization through dimerization. Furthermore, the ANM analysis also demonstrated the significant effects on protein dynamics due to protein-membrane couplings. Specifically, the incorporation of membrane constraints into ANM, as evidenced in the study of the glutamate transporter GltPh, revealed that the constraints imposed by the membrane selectively guide and facilitate the large-scale conformational transitions of the protein required for substrate transport. This approach highlighted that the presence of the membrane not only lowers the overall energy required for these transitions but also selectively modulates the motions of the transporter, thereby providing a more comprehensive understanding of the protein's dynamics in its native environment.[34]

How does a simple linear model provide a reasonable description of conformational behavior?

It is important to note that the ANM predicts movements within an energy well, by definition. So, one may wonder how it is possible to achieve a high correlation between ANM-predicted modes of motion and the conformational transitions observed experimentally between two substates presumably separated by an energy barrier. The answer lies in three facts: first, coarse-graining smooths out the energy surface such that substates separated by a low energy barrier (which are often required for transient events such as alternating access here) could be even merged within a shallow global minimum. The slowest modes of motion, which by definition refer to the directions of motions that require the least energy ascent near the original energy minimum, may as well point to such easily accessible substates that may be located within the same global minimum in the smoothed-out surface. This is the first advantage of coarse-graining, that of eliminating local barriers which might otherwise form traps on the energy landscape. Secondly, the ENMs are mathematically exact, i.e. they provide a unique analytical solution for each structure defined by its complete topology of inter-residue contacts. The loss of physicochemical accuracy is therefore partly compensated by the mathematical rigor of the network analysis, which helps identify the most cooperative directions of reconfiguration uniquely favored by the overall architecture. Thirdly, it is worth noting that these directions are by definition those favored by the conformational entropy of the overall architecture, which calls attention to the significant driving force of entropy maximization in dominating the movements accessible for achieving function.

Yet, the lack of atomic details, and the need to consider residue specificity, especially near active regions of proteins or at drug-binding pockets has been a drawback for the utility of ENMs in specific applications, which led to the development of hybrid approaches as described next.

Enhancing conformational sampling in all-atom MD simulations and docking simulations

ENMs predict low-frequency movements but do not capture events at atomic resolution because the system is coarse-grained to N nodes (interaction units) from the M total number of atoms, and the use of uniform harmonic potentials is a major simplification that overlooks residue specificity. Consequently, the degrees of freedom, and thus the number of modes, decrease from $3M\text{-}6$ *to* $3N\text{-}6$, and all non-linear interactions are neglected. For this reason, all-atom details are absent in the ENM predictions. As described previously, the global motions predicted by ENMs show good agreement with those sampled through computationally highly demanding extensive MD simulations,[4] which have $3M$ degrees of freedom, and this is partly due to the dominant effect of entropy and the fact that ENMs take rigorous account of overall connectivity or topology of contacts.

Recent advances in computational software and hardware have expanded the standard system size and simulation lengths for all-atom MD simulations to roughly 100,000 atoms and several hundred nanoseconds, respectively. However, for system sizes of about 1 million atoms and larger, the number of reported all-atom MD studies

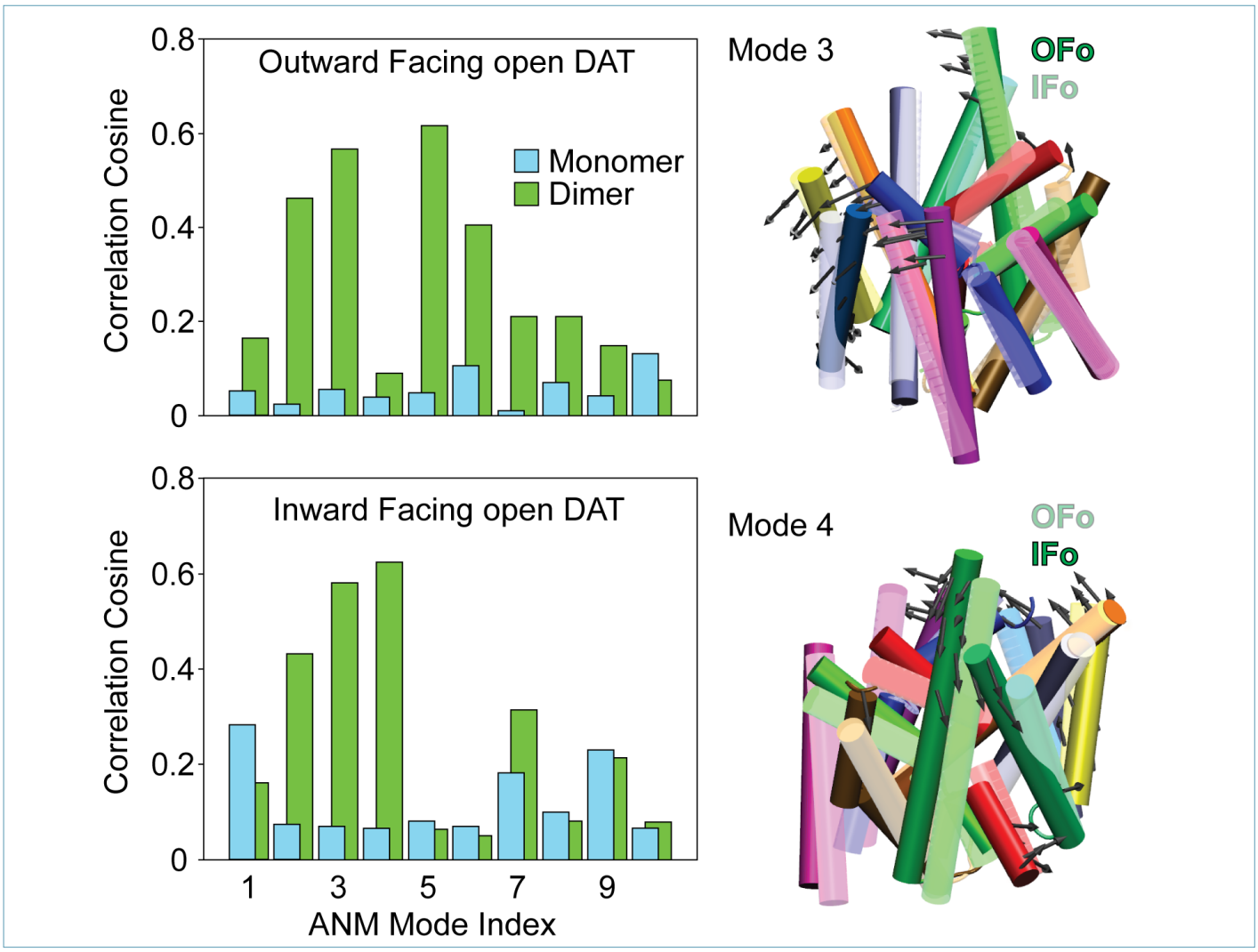

Figure 3: The ability of ANM modes to predict the IF ↔ OF transition for monomeric and dimeric dopamine transporter (DAT). The bars represent the correlation cosines, $cos(d, u_k)$, between the motions driven by the individual modes (10 slowest modes, $1 \leq k \leq 10$, along the abscissa) and the deformation vector $\boldsymbol{d} = \boldsymbol{R}(\mathrm{IFo}) - \boldsymbol{R}(\mathrm{OFo})$ experimentally observed between the IF open and OF open states of DAT in either direction, starting from (A) OF open and (B) IF open structures. Results for the protomers in the DAT dimer and the DAT monomers are indicated by the respective green and light-blue bars. The ribbon diagrams illustrate the collective motions driven by ANM mode 3 (A) and mode 4 (B) that achieve the highest correlation cosine with experiments. DAT monomers used as input were taken from MD simulations[35] and dimers were built[29] using them. For further details, please refer to Gur et al.[29]

are scarce, and the simulation times, often limited to microseconds, can become insufficient. Owing to the correlation between ENM-predicted global modes and protein functional dynamics, hybrid MD simulation techniques[36,37] that incorporate ENM modes to enhance conformational sampling, have become increasingly popular. ENM-MD hybrid methods employ ENM-predicted global modes to depict large-scale motions, such as domain rearrangements, while delivering atomic coordinates and taking account of detailed residue-specific energetics via all-atom MD simulations. Hybrid methods facilitate a natural selection of reaction coordinates guided by the ENM modes intrinsically accessible to the protein fold, while also incorporating non-linear effects that are absent in ENM calculations. In fact, a recent critical evaluation of hybrid methods based on integrated ENM and MD simulations demonstrated enhanced computational efficiency across three different techniques without sacrificing accuracy.[38] These methods initially focused on sampling enhancement, but their extensions have also proven effective in delivering accurate predictions of the free energy variations and surfaces related to functional protein motions. For example, the Collective MD (coMD)[39] methodology harnesses collective modes of motion encoded by the fold and evaluates interactions and energetics through MD simulations. The core strategy involves selecting ANM-predicted modes using a Monte Carlo/Metropolis algorithm, collectively deforming the structure along these modes, and directing all-atom MD simulations to sample collective changes in the overall structure at an atomic resolution. In the case of the Leucine transporter (LeuT), a bacterial membrane protein and an important ortholog of neurotransmitter transporters in humans as mentioned above, coMD successfully enabled the sampling of the conformational transition between its extreme endpoints, the OFo and IFo structures at full atomic resolution.[40] Subsequently, by initiating unbiased simulations from intermediate conformations across the transition, the free energy surface of LeuT's functional dynamics was generated (Figure 4).[40]

Hybrid methods have also found utility in the incorporation of receptor flexibility in docking studies. For example, ClustENMD approach combines deformation of structures along ANM modes, clustering and relaxation simulations using standard MD[41]. The utility of this method in ensemble

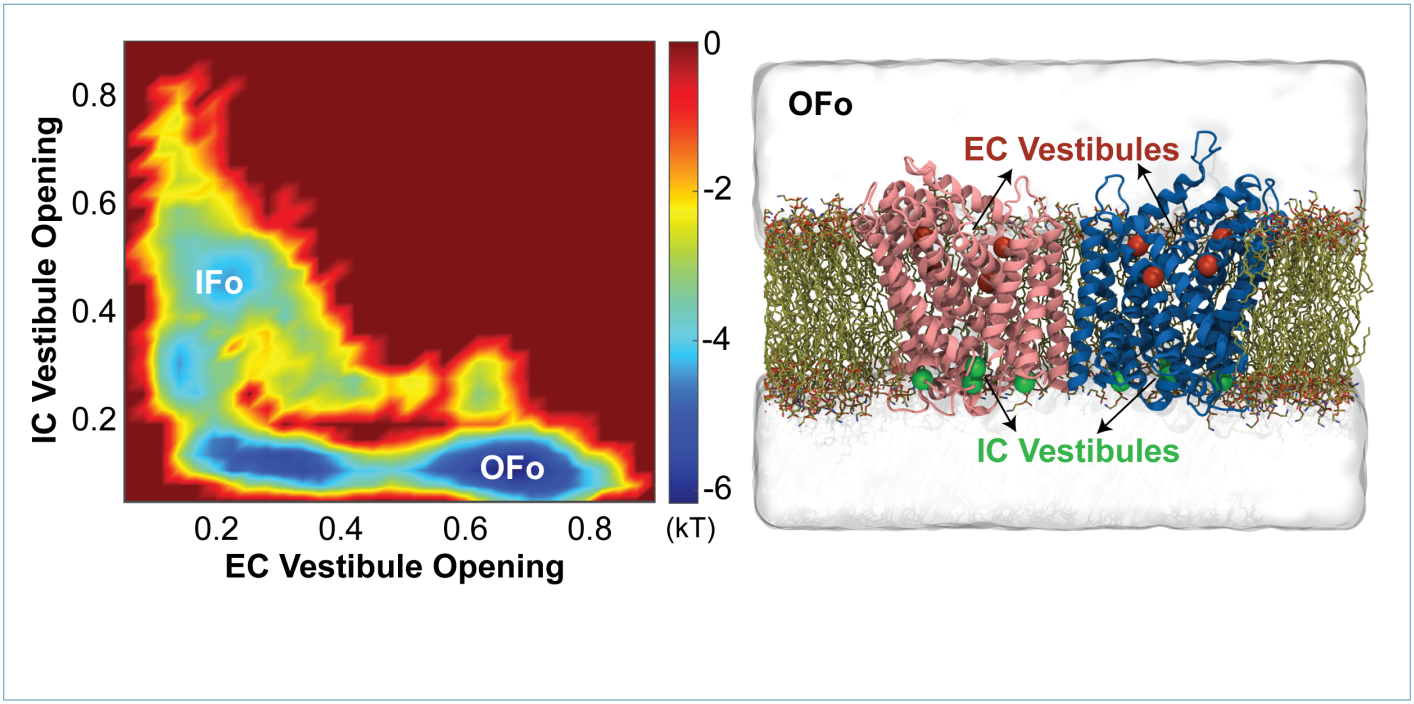

Figure 4: Free energy surface of LeuT obtained through the combined use of ENM and MD simulations via coMD. *Left panel:* **The free energy surface is shown as a function of the degree of EC and IC vestibules' openings. Refer to Gur et al.[40] for details.** *Right panel:* **Molecular structure of the outward-facing (OFo) LeuT dimer in the presence of membrane and solvent. Residues used to define EC and IC vestibules are shown in red and green respectively.**

docking lies in its ability to yield a spectrum of conformations for the docking target. This has been demonstrated in predicting ligand binding to highly flexible receptors[42] as well as protein-protein/DNA associations[43].

Combining strength of ENMs and Cryo-Electron Microscopy in the era of cryo-EM resolution revolution

Cryo-electron microscopy (cryo-EM) is a technique that has gained increased importance in recent years, as it now allows for the determination of large biomolecular structures at near-atomic resolution by flash-freezing samples and capturing images using an electron microscope. This method generates high-resolution 3D structures of macromolecular complexes, providing detailed structural information on alternative substates, but often lacking dynamic and temporal insights into the mechanisms of passages between those substates. To combine the strengths of cryo-EM structure determination and ENM analysis of collective dynamics, various hybrid methodologies have been developed.[37] One of them is the Hybrid Electron Microscopy Normal Mode Analysis (HEMNMA)[44], which integrates NMA with cryo-EM data. HEMNMA uses the global motions predicted by ENMs to deform structures into conformations that fit the cryo-EM density maps. This approach enables the extraction of continuous conformational changes from cryo-EM data, providing a comprehensive view of the structural dynamics. Another approach is the combination of MD simulations with ENMs, such as the MDeNM-EMfit[45] method. This technique involves performing MD simulations driven by ENM-predicted normal modes describing collective motions, while refining local elements through full atomic force field representations. This hybrid method enhances the sampling efficiency and accuracy of conformational landscapes derived from cryo-EM data by progressively directing the sampling toward conformations with higher correlations with the experimental cryo-EM maps.

Conclusion

The journey of ENMs from theoretical constructs to indispensable tools in computational biology underscores their profound impact on the study of protein dynamics. ENMs, with their simplified harmonic approximations, provide a robust framework for understanding the intrinsic motions and flexibility of proteins. The ENM-based application programming interface ProDy[46] has met an important need in the structural and computational biology community, evidenced by its broad usage worldwide. This review has highlighted the significant milestones in the development of ENMs, from the GNM to the ANM, their applications in predicting experimental observations and their utility in their combined use with all-atom MD simulations in hybrid simulations. The integration of ENMs with MD simulations and cryo-EM is currently a broadly pursued area toward enabling detailed insights into protein conformational landscapes and functional mechanisms. As computational and experimental techniques continue to evolve, ENMs will undoubtedly play a crucial role in advancing our understanding of complex biological systems. Their ability to bridge the gap between computational efficiency and biological accuracy makes them key in the ongoing quest to solve the molecular details of life. The future of ENMs lies in their continued refinement and application across diverse biological contexts, promising even greater contributions to the field of structural biology, in tandem with advances in experimental characterization of structures or contact topologies at multiple scales. A recent example along such directions is the GNM-based evaluation of chromosomal dynamics based on contact between gene loci collected by Hi-C technology for entire genomes, allowing for the first time a structure-based understanding of cell-specific gene expression and gene-gene correlations patterns for the entire chromatin.[47-49]

References

1 Dill, K. A. & Chan, H. S. From Levinthal to pathways to funnels. *Nat Struct Biol* **4**, 10-19 (1997). https://doi.org:10.1038/nsb0197-10

2 Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z. & Socci, N. D. Toward an outline of the topography of a realistic protein-folding funnel. *Proc Natl Acad Sci U S A* **92**, 3626-3630 (1995). https://doi.org:10.1073/pnas.92.8.3626

3 Hayward, S., Kitao, A. & Go, N. Harmonicity and anharmonicity in protein dynamics: a normal mode analysis and principal component analysis. *Proteins* **23**, 177-186 (1995). https://doi.org:10.1002/prot.340230207

4 Gur, M., Zomot, E. & Bahar, I. Global motions exhibited by proteins in micro- to milliseconds simulations concur with anisotropic network model predictions. *J Chem Phys* **139**, 121912 (2013). https://doi.org:10.1063/1.4816375

5 Straub, J. E. & Thirumalai, D. Exploring the energy landscape in proteins. *Proc Natl Acad Sci U S A* **90**, 809-813 (1993). https://doi.org:10.1073/pnas.90.3.809

6 Cui, Q. & Bahar, I. *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems (1st ed.).* (Chapman and Hall/CRC, 2005).

7 Go, N., Noguti, T. & Nishikawa, T. Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc Natl Acad Sci U S A* **80**, 3696-3700 (1983). https://doi.org:10.1073/pnas.80.12.3696

8 Brooks, B. & Karplus, M. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc Natl Acad Sci U S A* **80**, 6571-6575 (1983). https://doi.org:10.1073/pnas.80.21.6571

9 Levitt, M., Sander, C. & Stern, P. S. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *J Mol Biol* **181**, 423-447 (1985). https://doi.org:10.1016/0022-2836(85)90230-x

10 Golcuk, M., Yilmaz, S. Z., Yildiz, A. & Gur, M. The mechanism and energetics of the dynein priming stroke. *Structure* **32**, 603-610.e604 (2024). https://doi.org:10.1016/j.str.2024.02.003

11 Tirion, M. M. Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. *Phys Rev Lett* **77**, 1905-1908 (1996). https://doi.org:10.1103/PhysRevLett.77.1905

12 Bahar, I., Atilgan, A. R. & Erman, B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold Des* **2**, 173-181 (1997). https://doi.org:10.1016/s1359-0278(97)00024-2

13 Flory, P. J. & Erman, B. Theory of elasticity of polymer networks. 3. *Macromolecules* **15**, 800-806 (1982). https://doi.org:10.1021/ma00231a022

14 Kloczkowski, A., Mark, J. E. & Erman, B. Chain dimensions and fluctuations in random elastomeric networks. 1. Phantom Gaussian networks in the undeformed state. *Macromolecules* **22**, 1423-1432 (1989). https://doi.org:10.1021/ma00193a070

15 Bahar, I. & Jernigan, R. L. Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. *J Mol Biol* **266**, 195-214 (1997). https://doi.org:10.1006/jmbi.1996.0758

16 Haliloglu, T., Bahar, I. & Erman, B. Gaussian Dynamics of Folded Proteins. *Physical Review Letters* **79**, 3090-3093 (1997). https://doi.org:10.1103/PhysRevLett.79.3090

17 Bahar, I., Atilgan, A. R., Demirel, M. C. & Erman, B. Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability. *Physical review letters* **80**, 2733 (1998). https://doi.org:10.1103/PhysRevLett.80.2733

18 Bahar, I., Lezon, T. R., Yang, L. W. & Eyal, E. Global dynamics of proteins: bridging between structure and function. *Annu Rev Biophys* **39**, 23-42 (2010). https://doi.org:10.1146/annurev.biophys.093008.131258

19 Zhang, Y. *et al.* Intrinsic dynamics is evolutionarily optimized to enable allosteric behavior. *Curr Opin Struct Biol* **62**, 14-21 (2020). https://doi.org:10.1016/j.sbi.2019.11.002

20 Bahar, I., Lezon, T. R., Bakan, A. & Shrivastava, I. H. Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. *Chem Rev* **110**, 1463-1497 (2010). https://doi.org:10.1021/cr900095e

21 Atilgan, A. R. *et al.* Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J* **80**, 505-515 (2001). https://doi.org:10.1016/s0006-3495(01)76033-x

22 Delarue, M. & Sanejouand, Y. H. Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model. *J Mol Biol* **320**, 1011-1024 (2002). https://doi.org:10.1016/s0022-2836(02)00562-4

23 Xia, F., Tong, D. & Lu, L. Robust Heterogeneous Anisotropic Elastic Network Model Precisely Reproduces the Experimental B-factors of Biomolecules. *J Chem Theory Comput* **9**, 3704-3714 (2013). https://doi.org:10.1021/ct4002575

24 Erman, B. The gaussian network model: precise prediction of residue fluctuations and application to binding problems. *Biophys J* **91**, 3589-3599 (2006). https://doi.org:10.1529/biophysj.106.090803

25 Eyal, E., Yang, L. W. & Bahar, I. Anisotropic network model: systematic evaluation and a new web interface. *Bioinformatics* **22**, 2619-2627 (2006). https://doi.org:10.1093/bioinformatics/btl448

26 Doruker, P., Atilgan, A. R. & Bahar, I. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to alpha-amylase inhibitor. *Proteins* **40**, 512-524 (2000). https://doi.org:10.1002/1097-0134(20000815)40:3<512::AID-PROT180>3.0.CO;2-M

27 Leioatts, N., Romo, T. D. & Grossfield, A. Elastic Network Models are Robust to Variations in Formalism. *J Chem Theory Comput* **8**, 2424-2434 (2012). https://doi.org:10.1021/ct3000316

28 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 U S A* **106**, 14349-14354 (2009). https://doi.org:10.1073/pnas.0904214106

29 Gur, M., Cheng, M. H., Zomot, E. & Bahar, I. Effect of Dimerization on the Dynamics of Neurotransmitter:Sodium Symporters. *J Phys Chem B* **121**, 3657-3666 (2017). https://doi.org:10.1021/acs.jpcb.6b09876

30 Kaynak, B. T. *et al.* Cooperative mechanics of PR65 scaffold underlies the allosteric regulation of the phosphatase PP2A. *Structure* **31**, 607-618.e603 (2023). https://doi.org:10.1016/j.str.2023.02.012

31 Banerjee, A. *et al.* Influence of point mutations on PR65 conformational adaptability: Insights from molecular simulations and nanoaperture optical tweezers. *Sci Adv* **10**, eadn2208 (2024). https://doi.org:10.1126/sciadv.adn2208

32 Groves, M. R., Hanlon, N., Turowski, P., Hemmings, B. A. & Barford, D. The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell* **96**, 99-110 (1999). https://doi.org:10.1016/s0092-8674(00)80963-0

33 Leonard, D. *et al.* Selective PP2A Enhancement through Biased Heterotrimer Stabilization. *Cell* **181**, 688-701.e616 (2020). https://doi.org:10.1016/j.cell.2020.03.038

34 Lezon, T. R. & Bahar, I. Constraints imposed by the membrane selectively guide the alternating access dynamics of the glutamate transporter GltPh. *Biophys J* **102**, 1331-1340 (2012). https://doi.org:10.1016/j.bpj.2012.02.028

35 Cheng, M. H. & Bahar, I. Molecular Mechanism of Dopamine Transport by Human Dopamine Transporter. *Structure* **23**, 2171-2181 (2015). https://doi.org:10.1016/j.str.2015.09.001

36 Krieger, J. M., Doruker, P., Scott, A. L., Perahia, D. & Bahar, I. Towards gaining sight of multiscale events: utilizing network models and normal modes in hybrid methods. *Curr Opin Struct Biol* **64**, 34-41 (2020). https://doi.org:10.1016/j.sbi.2020.05.013

37 Costa, M. G. S., Gur, M., Krieger, J. M. & Bahar, I. Computational biophysics meets cryo-EM revolution in the search for the functional dynamics of biomolecular systems. *WIREs Computational Molecular Science* **14**, e1689 (2024). https://doi.org:https://doi.org/10.1002/wcms.1689

38 Kaynak, B. T. *et al.* Sampling of Protein Conformational Space Using Hybrid Simulations: A Critical Assessment of Recent Methods. *Frontiers in molecular biosciences* **9**, 832847 (2022). https://doi.org:10.3389/fmolb.2022.832847

39 Gur, M., Madura, J. D. & Bahar, I. Global transitions of proteins explored by a multiscale hybrid methodology: application to adenylate kinase. *Biophys J* **105**, 1643-1652 (2013). https://doi.org:10.1016/j.bpj.2013.07.058

40 Gur, M., Zomot, E., Cheng, M. H. & Bahar, I. Energy landscape of LeuT from molecular simulations. *J Chem Phys* **143**, 243134 (2015). https://doi.org:10.1063/1.4936133

41 Kaynak, B. T., Zhang, S., Bahar, I. & Doruker, P. ClustENMD: efficient sampling of biomolecular conformational space at atomic resolution. *Bioinformatics (Oxford, England)* **37**, 3956-3958 (2021). https://doi.org:10.1093/bioinformatics/btab496

42 Kurkcuoglu, Z. & Doruker, P. Ligand Docking to Intermediate and Close-To-Bound Conformers Generated by an Elastic Network Model Based Algorithm for Highly Flexible Proteins. *PLoS One* **11**, e0158063 (2016). https://doi.org:10.1371/journal.pone.0158063

43 Kurkcuoglu, Z. & Bonvin, A. Pre- and post-docking sampling of conformational changes using ClustENM and HADDOCK for protein-protein and protein-DNA systems. *Proteins* **88**, 292-306 (2020). https://doi.org:10.1002/prot.25802

44 Jin, Q. *et al.* Iterative elastic 3D-to-2D alignment method using normal modes for studying structural dynamics of large macromolecular complexes. *Structure* **22**, 496-506 (2014). https://doi.org:10.1016/j.str.2014.01.004

45 Costa, M. G. S., Fagnen, C., Vénien-Bryan, C. & Perahia, D. A New Strategy for Atomic Flexible Fitting in Cryo-EM Maps by Molecular Dynamics with Excited Normal Modes (MDeNM-EMfit). *J Chem Inf Model* **60**, 2419-2423 (2020). https://doi.org:10.1021/acs.jcim.9b01148

46 Zhang, S. *et al.* ProDy 2.0: increased scale and scope after 10 years of protein dynamics modelling with Python. *Bioinformatics* **37**, 3657-3659 (2021). https://doi.org:10.1093/bioinformatics/btab187

47 Zhang, S., Chen, F. & Bahar, I. Differences in the intrinsic spatial dynamics of the chromatin contribute to cell differentiation. *Nucleic Acids Res* **48**, 1131-1145 (2020). https://doi.org:10.1093/nar/gkz1102

48 Banerjee, A., Zhang, S. & Bahar, I. Genome structural dynamics: insights from Gaussian network analysis of Hi-C data. *Brief Funct Genomics* (2024). https://doi.org:10.1093/bfgp/elae014

49 Sauerwald, N., Zhang, S., Kingsford, C. & Bahar, I. Chromosomal dynamics predicted by an elastic network model explains genome-wide accessibility and long-range couplings. *Nucleic Acids Res* **45**, 3663-3673 (2017). https://doi.org:10.1093/nar/gkx172