Published in final edited form as:
Chem Rev. 2010 March 10; 110(3): 1463–1497. doi:10.1021/cr900095e.

Normal Mode Analysis of Biomolecular Structures: Functional Mechanisms of Membrane Proteins

Ivet Bahar[*], **Timothy R. Lezon**[†], **Ahmet Bakan**[†], and **Indira H. Shrivastava**
Department of Computational Biology, School of Medicine, University of Pittsburgh, 3064 BST3, 3501 Fifth Avenue, Pittsburgh, Pennsylvania 15213

1. Introduction

1.1. Protein Dynamics and Allostery

1.1.1. Dynamic Equilibrium between Pre-existing Conformations—The ability of macromolecules to sample *an ensemble of conformations* has been evident for decades, starting from the statistical mechanical theory and simulations of polymers.[1–3] A polymer chain of N atoms enjoys $3N - 6$ internal degrees of freedom, which gives rise to infinitely many conformations. Even a simple model of $N = 100$ atoms where bond lengths and bond angles are fixed, and dihedral angles are restricted to discrete isomeric states—say three states per bond—has access to $3^{N-3} = 1.9 \times 10^{46}$ conformations. Proteins, too, are polymers, and have access to ensembles of conformations. The main structural difference between proteins and other chain molecules is that, under physiological conditions, proteins sample a significantly narrower distribution of conformations compared to disordered polymers. Their conformational variations are confined to the neighborhood of a global energy minimum that defines their "native state".

While the native state has been traditionally viewed as a "unique structure" selected or encoded by the particular amino acid sequence, it is now established by theory, computations, and experiments, after the work of pioneering scientists in the field,[4–15] that the native state actually represents an ensemble of *microstates*: these microstates maintain the overall "fold" and usually share common secondary structure, but they differ in their detailed atomic coordinates. Differences are manifested by variations in bond lengths, bond angles, dihedral angles, loop conformations, substructure packing, or even entire domain or subunit positions and orientations.

Importantly, these microstates are not static: there is a *dynamic equilibrium* which allows for continual interconversions between them while maintaining their probability distribution. These "jigglings and wigglings of atoms" as expressed by Feynman,[16] and clearly observed in molecular dynamics (MD) simulations, were originally viewed as random events, or stochastic properties, hardly relevant to biological function. They essentially account for local relaxation phenomena in the nanoseconds regime, which may facilitate, for example, the diffusion of oxygen into the heme cavity of myoglobin[17] or the permeation of ions across selectivity filters in ion channels.[18–20] However, recent studies indicate that these thermal fluctuations may not only passively facilitate but also actively drive concerted domain movements and/or allosteric interactions, such as those required for substrate binding, ion

* To whom correspondence should be addressed. Phone: 412 648 3332. Fax: 412 648 3163. bahar@pitt.edu..
†These authors contributed equally to this work.

channel gating, or catalytic function.[15,21–34] Figure 1 provides an overview of the broad range of equilibrium motions accessible under native state conditions, ranging from bond length vibrations, of the order of femtoseconds, to coupled movements of multimeric substructures, of the order of milliseconds or seconds.

1.1.2. Functional Significance of Collective Motions—In the last two decades, there has been a surge in the number of studies based on *principal components analysis* (PCA)[36] of biomolecular structures and dynamics. These studies have proven useful in unraveling the *collective modes*, and in particular those at the low frequency end of the mode spectrum, that underlie the equilibrium dynamics of proteins.37 Normal mode analysis (NMA) of equilibrium structures,[38,39] essential dynamics analysis (EDA) of the covariance matrices retrieved from MD runs,40 and singular value decomposition (SVD) of MD or Monte Carlo (MC) trajectories[41–43] all fall in this category of PCA-based methods. Recently, a server has been developed to efficiently perform such calculations using a variety of input structures.44

PCA-based studies provide increasing support to the view that the apparently random fluctuations of proteins under native state conditions conceal contributions from highly cooperative movements (e.g., concerted opening and closing of domains) that are directly relevant to biological function. Functional movements indeed involve passages between collections of *microstates* or substates that coexist in a dynamic equilibrium (Figure 2). The most cooperative motions usually occur at the low frequency end of the mode spectrum. These modes engage large substructures, if not the entire structure, hence their designation as *global* or *essential* modes. They are *intrinsically accessible* to biomolecules, arising solely from structure. In a sense, in the same way as sequence encodes structure, structure encodes the equilibrium dynamics. We refer to these global movements that are collectively encoded by the 3-dimensional (3D) structure as *intrinsic motions* of the examined protein, intrinsic to the protein fold or topology of native contacts. Biomolecular structures conceivably evolved to favor the global modes that help them achieve their biological or allosteric functions.[21] Briefly, the emerging paradigm is *structure-encodes-dynamics-encodes-function*, and an evolutionary pressure originating from functional dynamics requirements may have selected the relatively small space of functional structures.

The predisposition of proteins to undergo functional changes in structure is now supported by numerous experimental and computational studies, and an increasing amount of data demonstrates that allosteric responses are driven by intrinsically accessible motions.[15,23,24,45–51] These studies have brought a new understanding to the role of collective dynamics in protein functions, demonstrating in particular how the functions of membrane proteins such as signal transduction, pore opening, ion gating, or substrate translocation are enabled by the cooperative movements of symmetrically arranged subunits. These findings are in support of the original Monod–Wyman–Changeux (MWC) view of allosteric effects,[52,53] the main tenets of which are *predisposition* of the structure to access alternative conformations via cooperative changes in structure (simultaneously engaging all subunits) and *selection* from this pool of accessible conformation to achieve biological function in the presence of ligand/substrate binding. Recent findings on the relevance of global modes to functional dynamics are presented below for select, widely studied membrane proteins. The goal here is to review NMA-based computational methods and their applications to membrane proteins. We will also discuss recent developments for improving the methodology and its implementation in structure refinement and drug discovery methods.

1.1.3. Normal Mode Analysis: An Old Technique That Recently Found a Revival in Molecular Biology—Normal mode analysis provides information on the equilibrium modes accessible to a system, assuming that the system is stabilized by harmonic potentials. It has been used for several decades in studying classical physical phenomena such as atomic

vibrational spectra and transport in the solid state. Its application to proteins dates back to the early 1980s.[54–57] However, only in the past decade has it become a tool widely used for exploring functional motions. A major reason behind its broader use is the observation that global modes elucidated by NMA bear functional significance. This feature became even more evident with the use of simplified models in coarse-grained (CG) NMA.[38]

From a physical perspective, the global modes simply represent reconfigurations along directions (principal axes) that are most easily accessible (require the least energy ascent for a given size deformation) on the multidimensional energy landscape. Mode frequency (squared) provides a measure of the curvature (or *stiffness*) of the energy landscape along a given mode direction, with lower frequency modes being softer motions. Given that the energy landscape is, in turn, defined by the molecular structure, the global modes are structure-encoded by definition.

A striking feature of NMA and other PCAs of biomolecular structures is the observed *robustness* of the global modes to details in atomic coordinates or specific interatomic interactions. The global modes are defined by the entire structure (or architecture). Their insensitivity to local interactions, or to the specific energy functions and parameters that define the force field, presumably results from their systemic nature. As evidenced by the pioneering study of Tirion,[58] a hypothetical force field with uniform (single-parameter) harmonic potentials yields global modes almost indistinguishable from those obtained from a detailed force field with specific nonlinear terms. The property that apparently dominates the shape of the global modes is *the network of inter-residue contacts*, which is a purely geometric quantity defined by the overall shape, form, or native contact topology of the protein.[59,60]

1.1.4. Elastic Network Models Inspired by the Robustness of Global Modes—The insensitivity of global modes to structural and energetic details has now been confirmed in many studies[61–66] and led to the following question: If these modes are not sensitive to structural and energetic details, why not use CG models to elucidate such collective movements? This line of thought opened the way to the adoption of elastic network models (ENMs) for exploring protein dynamics.[38,39] ENMs take rigorous account of the topology of contacts. In this respect, they may be viewed as Go models which are also based on native contact maps.[4–6] The major difference is, however, that the network representation adopted in ENMs permits us to take advantage of methods of NMA or spectral graph theory to obtain analytical solutions for equilibrium dynamics that can be readily implemented in efficient computational algorithms. The main advantages of ENM-based approaches are indeed (i) their ability to provide an *exact* solution for the *unique* dynamics of each structure and (ii) their applicability to large biomolecular complexes and assemblies beyond the range of atomic simulations.

The first such simplified model, the Gaussian network model (GNM), was introduced a decade ago,[67,68] inspired by the work of Tirion.[58] GNM is based on the theory of elasticity set forth by Flory and co-workers[69–73] for polymer networks. The structure is represented as a network of nodes (α-carbons) and elastic springs. The springs connect the α-carbon pairs that are closer than a cutoff distance, R_c, in the native structure. A Kirchhoff matrix of inter-residue contacts, $\mathbf{\Gamma}$, is the sole determinant of equilibrium dynamics. The accessible spectrum of relaxation modes is computed using statistical mechanical theories of solid state physics and/or graph-theoretic methods that have found wide utility in other applications. The GNM has been shortly followed by other ENMs, including, in particular, Hinsen's model with distance-dependent force constants[74,75] and the anisotropic network model (ANM),[76–79] which will be described in some detail in section 2.3.2.

A major reason behind the broadening recognition of NMA as a tool for exploring functional motions of proteins is the observation that global modes elucidated by NMA bear functional significance. For example, the allosteric change in conformation undergone by hemoglobin from its tense (T) form to its relaxed (R) form has been shown by both full atomic NMA[80] and ANM31 to match closely the collective motions along the second lowest frequency mode intrinsically accessible to the original structure. The ratchet-like motion of the two subunits of the ribosome is enabled by the third slowest mode;81،82 or in general, the opening/closing of domains/subunits in many enzymes conforms to their low frequency modes.[32,79] It is now broadly recognized that ligand binding cooperatively triggers collective movements and stabilizes conformers that are already favored by, or accessible to, the unbound protein structure.[30]

In recent years, ENM-based NMAs have contributed to improving our understanding of the collective dynamics of membrane proteins, among other allosteric systems. Understanding the functional motions of membrane proteins is significant from both biological and pharmaceutical points of view, as described in section 1.2.1. Progress in this field has been slow, however, due to the scarcity of structural data and the complexity of the involved multiscale interactions. The majority of structure-based computations performed for membrane proteins in the past decade focused on localized events, such as passage of ions across a selectivity filter, which are observable by MD simulations of tens of nanoseconds. The computational study of events such as ion/substrate gating, on the other hand, has been limited by the more cooperative nature of associated changes in structure. Models and methods designed to observe longer time or larger size windows, not obscured by atomic details or random noise, are needed in this case. CG NMAs and their extensions and combinations with atomic simulations[44,83–88] are beginning to fill this gap. The applications summarized in section 3 permit us to observe for the first time a new regime of motions at residue-level detail, providing insight into mechanisms of pore opening, ion gating, or allosteric signal transduction by membrane proteins.

1.2. Structural Dynamics of Membrane Proteins: Significance and Challenges

1.2.1. Classification, Biological Role, and Pharmacological Importance—

Membrane proteins are classified into two broad groups: integral membrane proteins (IMPs) and peripheral membrane proteins. IMPs are permanently located at the membrane; peripheral membrane proteins are temporarily attached, either to IMPs or to the peripheral regions of the membrane. IMPs include channels, receptors, transporters, and enzymes, in addition to cell adhesion and energy transduction proteins. They are divided into two broad groups depending on the degree to which they span the lipid bilayer: transmembrane (TM) (or polytopic) and integral monotopic. Integral monotopic proteins are attached to the membrane from one side, while TM proteins are typically composed of three domains: extracellular (EC), intracellular/cytoplasmic (CP), and TM domains (Figure 3). The present review focuses on the equilibrium dynamics of selected TM proteins that have been explored by NMA-based approaches in recent years (section 3).

The continual flow of ions and metabolites across the membranes is essential for many of life's processes. The membrane acts as an insurmountable barrier for the passage of ions and/or solutes into or out of the cell under equilibrium conditions, thus maintaining a net voltage difference between the cell interior and exterior, known as the *resting membrane potential*. TM proteins maintain the equilibrium concentrations of ions/substrates in the EC and CP regions, elicit or regulate cell signaling and energy transduction processes, regulate cell volume, or secrete electrolytes.[89,90] In particular, ion pumps and ion exchangers “actively” transport ions: they pump ions against their gradient by coupling the “uphill” process to an energy source such as ATP hydrolysis or the “downhill” movement of an ion or substrate.

Likewise, carriers transport substrates, against their concentration gradient in many cases, assisted by proton or ion counter- or cotransport. Ion channels, on the contrary, are usually viewed to be “passive” membrane proteins: they allow for “downhill” permeation of ions and may exhibit very high conduction rates.[91] The electrochemical gradients built by ion pumps serve as a driving source for ion channels and other transporters.[89]

With their locations at cell boundaries, membrane proteins are involved not only in the transport of endogenous substrates/ions but also in drug uptake[92] and drug action. While approximately 30% of sequenced genes encode membrane proteins, the fraction of membrane proteins among drug targets has been estimated to be 70% in 2001.[93] An updated distribution of drug targets is presented in Figure 4. The pie chart refers to 965 U.S. Food and Drug Administration (FDA) approved small-molecule drugs, obtained from DrugBank (http://www.drugbank.ca)[94] as primary source and refined using Therapeutic Target Database (DB),[95] Super-Target DB,96 and the literature.97 A total of 380 proteins are targeted by these drugs, most of which belong to the human genome. The corresponding molecular functions, retrieved from the PANTHER Classification System,98 are grouped into 71 functional categories. Figure 4 displays the most frequently targeted ten such categories. The top-ranking four categories are G-protein coupled receptors (GPCRs), nuclear hormone receptors, ligand-gated ion channels, and voltage-gated ion channels. These constitute targets for more than half of the drugs. These results are consistent with those recently compiled by Hopkins and co-workers,[99] apart from minor differences, presumably due to differences in the data set.

The membrane proteins that are most frequently targeted by small molecule drugs are histamine H1 receptors, α1-adrenergic receptors, and D2 dopamine receptors. All three are members of the GPCR family of proteins. These are succeeded by γ-aminobutyric-acid (GABA) A receptor α1, a ligand-gated ion channel. These proteins are still being investigated in relation to a broad spectrum of diseases including central nervous system disorders, allergies, inflammation, respiratory disorders, headache, and sleep disorders.100 Notably, most of the drugs currently in use were not initially developed to interact with a specific target protein but to induce certain phenotypes in cultured cell or animal assays.101 The identification of the targets followed the completion of the drug discovery process (a trial-and-error process using combinatorial chemistry rules). The importance of assessing drug targets and understanding the mechanistic aspects of drug–target associations became clear only in recent years. Knowledge of structure and dynamics of target proteins is now recognized to be a crucial element in making progress in rational drug discovery.102

1.2.2. Increasing Structural Data on Membrane Proteins—Although atomic resolution crystal structures of soluble proteins have historically been reported with an exponentially increasing frequency, similar progress has not been made for membrane proteins. Determination of membrane protein structures is difficult for a number of reasons. Their crystallization requires disruption of the bilayer, usually with detergents, which renders many of the physical methods of crystallization difficult or impractical. The amphipathic nature of membrane proteins and their inherent conformational flexibility also poses a problem for crystallization.[103] Yet another reason for the limited structural data on membrane proteins is their low concentrations in tissues, i.e., the difficulties in overexpressing and purifying them at the milligram level.

In spite of these limitations, important progress has been made in recent years in determining membrane protein structures.[104] With the advances in high-throughput techniques in structural biology, multiple protein targets are being cloned, expressed, and purified in parallel, with clones being generated on multiwell plates and crystallization trials being completed at the rate of 100 plates per day.[105,106] These techniques are now being advantageously used in membrane protein structure determination. Membrane proteins are unstable in detergent micelles, and

finding conditions that stabilize them helps in protein crystallization. Often, a series of detergents are tested, and the one that extracts the maximum quantity of soluble, active, homogeneous protein is used. Dodecyl maltoside is a detergent that provides such conditions.[107] The recently solved full-length KcsA structure is a nice example of engineering an enhanced stability at the C-terminal domain of the membrane protein by the use of synthetic antigen-binding fragments as crystallographic chaperones. We refer our readers to comprehensive reviews[103,104,108,109] for more information on advances in NMR spectroscopy,[108–116] X-ray crystallography,[117–122] electron crystallography of 2D crystals in the presence of lipids,[93] and infrared spectroscopy (IR).[123]

These advances resulted in a remarkable increase in the number of structurally known membrane proteins: while at the end of 2003, there were about 326 resolved membrane proteins (75 of them being unique), this number jumped to 859, including 182 unique structures by the end of 2008 (Figure 5). A comprehensive summary of membrane protein structures available to date is provided by the site (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html) maintained by Stephen White's lab. Access to all membrane protein structures in the Protein Data Bank (PDB) is provided by databases such as the PDB of Transmembrane Proteins[124] (http://pdbtm.enzim.hu/) and the Membrane PDB[125] (http://www.mpdb.ul.ie). The rapidly growing data in the PDB now holds promise for exploring the structure-based dynamics of membrane proteins.[91] The NMA-based studies presented in section 3 provide some examples of such explorations.

1.2.3. Multiscale Dynamics of Membrane Proteins Probed by Experiments—The biological function of many membrane proteins involves a transient change in structure, with the associated processes usually spanning a broad range of time scales (Figure 1). Since the early days of electrophysiology, theories and models have been developed to gain a better understanding of the structural basis of biological function and biochemical data.[89] Spectroscopic methods for examining conformational vibrations have been applied to proteins as early as 1952, long before the availability of detailed X-ray structures.[126] Their successful application to membrane proteins has been possible only after the development of sensitive instrumentation and sophisticated analysis techniques.[123]

Here, we briefly point to two spectroscopic methods, NMR and electron paramagnetic resonance (EPR), that emerge as powerful tools for probing protein dynamics over broad time scales, and we refer our readers to excellent reviews[127–136] for more information.

Structural dynamics probed by NMR: Many of the experimental data showing that structural dynamics play a central role in protein function currently originate from NMR spectroscopy.[22,131,135,137] While traditionally NMR studies have been limited to relatively small proteins (typically <50 kDa), the dynamics of significantly larger structures are now being explored with advances in labeling technology and solution NMR,[132] and motions in both short (pico- to nanoseconds) and long (micro- to milliseconds) time scales are being probed.[116,131,138–140] Using ^{13}C- and ^{2}H-NMR relaxation rates, Kay and co-workers showed that site-specific quantitative data could be collected for the correlation times of methyl groups for proteins of the order of 100 kDa;[141,142] for example, picoseconds-to-nanoseconds cross-correlation rates for intramethyl ^{1}H–^{1}H dipolar spin relaxation have been measured for a half proteasome complex.[133] Not all methyl sites undergo such fast dynamics, however. Those embedded in regions that undergo highly concerted subunit motions exhibit relaxation times of the order of milliseconds, as observed[143] in the α-rings of the 20S core particle proteasome —a molecular machine of 670 kDa. Finally, the ability of NMR experiments to separate local and global changes in conformation is noteworthy. A classical example is the T → R transition of aspartate transcarbamoylase,[144] the 2D ^{1}H–^{13}C spectra of which clearly evidenced the pre-existence of a dynamic equilibrium between the two forms and the stabilization of one (R)

upon ligand binding.[145] Residual dipolar couplings (RDCs) observed in NMR spectra are increasingly providing information on collective motions in the nano- to microsecond regime.[26,130,146] The PCA-based comparison of an ensemble of NMR models refined against RDCs[26] for ubiquitin with the X-ray structures of the same protein in the presence of different substrates showed (i) the accord between the conformations sampled in solution by the unbound protein and those assumed upon complexation with different substrates and (ii) the fact that conformational variations could be explained to a large extent by a few low frequency modes intrinsically accessible to the structures.

Structural dynamics probed by site-directed spin labeling (SDSL) combined with EPR[127–129,147]: In this technique, residues at selected sites are substituted, mostly by cysteines, followed by the selective modification of the sulfhydryl group with a nitroxide radical that serves as an EPR spin label/probe. A set of spin-labeled proteins is thus prepared, differing by the position of spin-labeled cysteines (Figure 6). The experiments allow for characterizing the mobilities of the spin-labeled residues and the changes in the distances between them. Applications of time-resolved SDSL-EPR to membrane proteins (e.g., to bacteriorhodopsin[148,149] and rhodopsin[134,150–155] by Hubbell, Khorana, and co-workers, to the K^+ channel from *Streptomyces lividans* (KcsA),[156–159] and to a prokaryotic mechanosensitive channel (MscL)[160] by Perozo and co-workers) have successfully provided information on interhelical movements or changes in tertiary contacts accompanying their functional rearrangements. Notably, fluctuations and correlations over wide time scales, from nanoseconds to milliseconds, can be examined by this technique, and changes in distance between pairs of labels separated by 20–60 Å can be probed by double electron resonance (DEER) with a resolution better than 1 Å.[161] A recent application to rhodopsin clearly demonstrated, for example, that the activation of the molecule is accompanied by an outward movement of TM helix 6 by about 5 Å.[161]

1.2.4. Structure-Based Models, Theory, and Computations—According to Moore's law, the cost of computing halves roughly every 1.5 to 2 years.[162] Looking back at MD simulations of biomolecules at atomic resolution, we have indeed progressed from tens of picoseconds in the early 1980s to tens of nanoseconds at present, roughly consistent with Moore's law. The progress in recent years in MD simulations of membrane proteins[163–166] has been enabled by advances in computing technology, algorithms, and methods.[167,168] Classical examples include the simulations of aquaporin,[33,169–172] gramicidin,[20,173–175] and KcsA.[18,176–179] Yet, the time scales of tens of nanoseconds, or even tenths of microseconds,[180] being accessed in advanced MD simulations still fall short of providing an accurate sampling of the complete conformational space that many multimeric membrane proteins explore under physiological conditions. Indeed, multimeric structure is recognized as essential for enabling highly cooperative structural rearrangements.[52] Physics-based CG models such as ENMs emerge as *approximate* structural models that provide *analytical* solutions for the collective dynamics of such complex systems, which cannot be determined with MD.

Here we present recent progress in exploring the dynamics of membrane proteins using NMA with ENM-based models. The proteins considered are grouped into three categories: ion channels, receptors, and transporters, presented in the respective sections 3.1–3.3, and among them we present NMA-based studies on gramicidin A (GA), potassium channels (KcsA and others), a MscL, the nicotinic acetylcholine receptor (nAChR), rhodopsin (as a prototypic GPCR), a glutamate transporter (Glt_{Ph}), and an ATP binding cassette (ABC) transporter (BtuCD) (Figure 7). Note that, in many cases, there is no clear-cut distinction between these categories; some receptors or transporters may also function as ion channels (e.g., nAChR is a ligand-gated ion channel, or glutamate transporter also functions as a chloride channel).

Membrane proteins are one of the most challenging biomolecular systems from theoretical and computational aspects, because their functional mechanisms involve not only the *global* dynamics of multiple subunits but also the highly specific *local* interactions (e.g., ionic interactions, solvation) in the presence of lipid bilayer and water molecules. The former group of motions is conveniently explored by CG models such as ENMs, but the latter requires detailed full atomic geometry and energy considerations and cannot be studied by ENMs. This broad range of events also spans a hierarchy of functionally significant time scales, which in turn requires adopting multiscale approaches. The combination of atomic-level resolution with high-level (low resolution) computational approaches is deemed as a promising path for interpreting experimental observations and determining structure–function relations.[20,181] Attempts at developing such integrated approaches by steering MD simulations along ENM directions85 or sampling transition pathways using adaptive ENM methods have recently been undertaken. These will be presented in section 4.2. Other recent studies support the utility of ENM-based approaches, not only for assessing functional dynamics but also for flexible docking of substrates and refining low resolution structures, and will be presented in section 4.2.

2. Theory

2.1. Principal Component Analysis of Experimentally Resolved Conformations

2.1.1. Definitions of System and Parameters—Principal component analysis of ensembles of structures is an orthogonal linear transformation that transforms data from the Cartesian coordinate system into a new system of collective coordinates.[36] The goal is to gain a simplified view of the structural variability in the examined data set by identifying the dominant directions of structural changes. The new coordinate system is such that the greatest variance in the data set lies along the first principal component (PC) axis, followed by the second PC axis, and so on. Here we will focus on the application of PCA methods to extract information on equilibrium dynamics. PCA is performed in this case for an ensemble of conformers such as the PDB structures for the same protein determined in the presence of different substrates, or NMR models for the same protein.[44] Alternatively, an ensemble of MD snapshots may be used. The utility of PCA for understanding functional dynamics is clearly demonstrated by a recent study of ubiquitin structures in which a single mode of motion was found to largely account for ubiquitin's recognition ability.26

Let us consider an ensemble of M conformations, for a protein of N interaction sites (N atoms or residues, or any CG representation of an interaction site). Each conformation, k, is described by a $3N$-dimensional vector consisting of the position vectors $\mathbf{R}_i^{(k)} = (x_i^{(k)}\, y_i^{(k)}\, z_i^{(k)})^{\mathrm{T}}$ of the N sites ($1 \le i \le N$) in that particular conformation, organized as

$$6\mathbf{q}^{(k)} = \left(\left(\mathbf{R}_1^{(k)}\right)^{\mathrm{T}}, \ldots, \left(\mathbf{R}_N^{(k)}\right)^{\mathrm{T}}\right)^{\mathbf{T}} = \left(x_1^{(k)}, y_1^{(k)}, \ldots, x_N^{(k)}, y_N^{(k)}, z_N^{(k)}\right)^{\mathbf{T}} \quad (1)$$

Likewise, we define a $3N$-dimensional fluctuation vector $\Delta\mathbf{q}^{(k)} = \mathbf{q}^{(k)} - \mathbf{q}^0$ for each conformation, describing the departure $\Delta\mathbf{R}_i^{(k)} = \mathbf{R}_i^{(k)} - \mathbf{R}_i^0$ in the position vectors of the N sites from their equilibrium position $\mathbf{R}_i^0 = (x_i^0\, y_i^0\, z_i^0)^{\mathbf{T}}$ (Figure 8). The equilibrium positions may be identified by the average over all snapshots from MD trajectories or over all optimally superimposed PDB structures. The PDB coordinates are usually assumed to be the equilibrium positions when performing NMA of a given structure using the ENM, as will be described below.

2.1.2. Covariance Matrix: A Measure of Correlations between Residue Motions

—In many applications, it is of interest to understand the type and strength of coupling between the variations in different degrees of freedom. The cross-correlations between the components of the fluctuations vectors are given by the averages $\langle \Delta q_i \Delta q_j \rangle$ over all conformations, which are conveniently organized in a $3N \times 3N$ covariance matrix **C**,

$$\mathbf{C} = M^{-1} \sum_k \Delta \mathbf{q}^{(k)} \Delta \mathbf{q}^{(k)\mathbf{T}} \tag{2}$$

A detailed description of equilibrium motions, including the mean-square fluctuations of individual sites and their cross-correlations, is provided by the covariance matrix **C**. The elements of **C** may alternatively be viewed as $N \times N$ blocks (or submatrices of size 3×3), $\mathbf{C}_{ij}$ each of the form

$$\mathbf{C}_{ij} = \begin{bmatrix} \langle (\Delta x_i \Delta x_j) \rangle & \langle (\Delta x_i \Delta y_j) \rangle & \langle (\Delta x_i \Delta z_j) \rangle \\ \langle (\Delta y_i \Delta x_j) \rangle & \langle (\Delta y_i \Delta y_j) \rangle & \langle (\Delta y_i \Delta z_j) \rangle \\ \langle (\Delta z_i \Delta x_j) \rangle & \langle (\Delta z_i \Delta y_j) \rangle & \langle (\Delta z_i \Delta z_j) \rangle \end{bmatrix} \tag{3}$$

Here we use boldface subscripts to designate a (sub)matrix or vector and lightface subscripts for scalars (e.g., elements of vectors or matrices). The sum of the diagonal elements of $\mathbf{C}_{ij}$,

$$\mathrm{tr}\{\mathbf{C}_{ij}\} = \langle \Delta x_i \Delta x_j \rangle + \langle \Delta y_i \Delta y_j \rangle + \langle \Delta z_i \Delta z_j \rangle = \langle \Delta \mathbf{R}_i \bullet \Delta \mathbf{R}_j \rangle \tag{4}$$

provides a measure of the cross-correlation between the fluctuations $\Delta \mathbf{R}_i$ and $\Delta \mathbf{R}_j$ of sites i and j; similarly, the mean-square fluctuations in the positions of individual sites are given by the trace of the diagonal submatrices, i.e., $\mathrm{tr}\{\mathbf{C}_{ii}\} = \langle (\Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_i) \rangle = \langle (\Delta \mathbf{R}_i)^2 \rangle$ using $i = j$ in eq 4. In many applications, it proves useful to analyze the $N \times N$ covariance matrix, $\bar{\mathbf{C}}$, composed of the correlations between the fluctuation vectors $\Delta \mathbf{R}_i$, themselves,

$$\bar{\mathbf{C}} = \begin{bmatrix} \langle (\Delta \mathbf{R}_1)^2 \rangle & \langle \Delta \mathbf{R}_1 \bullet \Delta \mathbf{R}_2 \rangle & \cdots & \cdots \\ \langle \Delta \mathbf{R}_2 \bullet \Delta \mathbf{R}_1 \rangle & \langle (\Delta \mathbf{R}_2)^2 \rangle & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \langle \Delta \mathbf{R}_N \bullet \Delta \mathbf{R}_1 \rangle & \cdots & \cdots & \langle (\Delta \mathbf{R}_N)^2 \rangle \end{bmatrix} \tag{5}$$

The fluctuations in the Cartesian space are mapped onto the space spanned by the $3N$ (or N) principal axes upon diagonalizing the covariance matrix **C** (or $\bar{\mathbf{C}}$) as

$$\mathbf{C} = \mathbf{P}\mathbf{S}\mathbf{P}^{\mathbf{T}} = \sum_{k=1}^{3N} \sigma_k \mathbf{p}_k \mathbf{p}_k^{\mathbf{T}} \tag{6}$$

where **P** is the unitary matrix of normalized displacements along the principal axes, also called the *principal modes* of structural changes, each given by a column $\mathbf{p}_k$, $(1 \le k \le 3N)$, and **S** is the diagonal matrix of eigenvalues $\sigma_1, \sigma_2, \ldots, \sigma_N$, usually ordered in descending order. The eigenvalues are directly proportional to the variance along the principal axes such that the fractional contribution of $\mathbf{p}_k$ to the structural variability in the data set is $f_k = \sigma_k / \sum_I \sigma_i$, where the summation is performed over all $3N$ modes. Equation 6 permits us to assess the contribution of each mode or subset of modes to the observed covariance. For example, the square

displacements in the position of the i^{th} interaction site induced by the top ranking m PC modes are

$$(\Delta \mathbf{R}_i)^2\Big|_{1\le k\le m} = \mathrm{tr}\left\{\left[\sum_{k=1}^{m}\sigma_k \mathbf{p}_k \mathbf{p}_k^{\mathrm{T}}\right]_{ii}\right\} \quad (7)$$

2.2. Normal Mode Analysis

2.2.1. Assumptions and Limitations—The mathematical theory of NMA is detailed in any classical mechanics text;[182] hence, here we will present only an outline of the theory as it pertains to its recent applications to proteins and their complexes. The essence of NMA is again the diagonalization of a $3N \times 3N$ matrix **H**, called the Hessian, the inverse of which yields the covariance matrix **C**.

The underlying assumption in NMA is that any given equilibrium system fluctuates about a single well-defined conformation and that the nature of these thermally induced fluctuations can be calculated assuming a simple harmonic form for the potential. This directly leads to a basic limitation of NMA: it is only valid in proximity to the equilibrium. As the system is displaced from its equilibrium conformation, the extent to which the harmonic approximation holds grows increasingly uncertain. The excursions from equilibrium along the normal modes must be closely monitored, lest one propose a conformational change in excess of the model's capabilities. In situations where the potential is calculated using exact units, for example when an atomistic force field is used as the kernel for the potential energy surface, then the magnitude of the excursions along the normal modes might be approximated by the absolute temperature of the system. When further coarse-graining is used, for example in the ENMs, then the absolute magnitudes of the interactions are unknown and even greater care must be used.

A second caveat to NMA is that the normal modes represent instantaneous displacements and are tangent to the direction of motion at equilibrium. The molecule often contains additional internal distance constraints that are not explicitly included in the NMA, such as fixed bond lengths or bond angles. All but the smallest distortions along the normal modes will violate these constraints unless measures are taken to satisfy them.[183,184] As a result, (i) NMA results are “accurate” in the immediate vicinity of the energy minimum, where the “direction” of motion is accurately predicted; large excursions in the conformational space may necessitate the use of an adaptive scheme to re-evaluate normal modes at a minimum, and (ii) NMA with ENMs is most useful for predicting the large-scale motions, which are insensitive to structural and energetic details; localized highly specific interactions, including in particular electrostatic interactions, cannot be precisely accounted for. Increasingly, NMA is used in conjunction with traditional MD or other simulation methods to explore large-scale motions in the presence of detailed atomic interactions.[85,183,185–189]

2.2.2. Underlying Potential and Hessian Matrix—For our purposes, the physical system under consideration is a molecular system containing N interaction sites, the Cartesian coordinates of which are given by eq 1. We will omit the superscript k here, since NMA is performed for a single structure ($M = 1$). Near the equilibrium conformation, $\mathbf{q}^0$, the potential energy can be expanded as a power series in **q** as

$$V(\mathbf{q}) = V(\mathbf{q}^0) + \sum_i \left(\frac{\partial V}{\partial q_i}\right)^0 \left(q_i - q_i^0\right) + \frac{1}{2}\sum_{i,j}\left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)^0 \left(q_i - q_i^0\right)\left(q_j - q_j^0\right) + \ldots \quad (8)$$

where superscripts of zero indicate the equilibrium conformation. The first term is the minimum value of the potential, which may be set to zero. The second term is identically zero at any local minimum of the potential. To second order, the potential is then a sum of pairwise potentials

$$\begin{aligned} V(\mathbf{q}) &= \frac{1}{2}\sum_{ij}\left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)^0 \left(q_i - q_i^0\right)\left(q_j - q_j^0\right) \\ &= \frac{1}{2}\sum_{i,j}\left(q_i - q_i^0\right) H_{ij}\left(q_j - q_j^0\right) = \frac{1}{2}\Delta\mathbf{q}^{\mathbf{T}}\mathbf{H}\Delta\mathbf{q} \end{aligned} \quad (9)$$

where **H** is the Hessian matrix obtained from the second derivatives of the potential with respect to the components of **q** (or Δ**q**):

$$H_{ij} = \left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)^0 \quad (10)$$

In the same way as the covariance matrix **C** is organized, **H** may be thought of as an $N \times N$ matrix of 3×3 submatrices, each of which describes the energetic contribution from the interaction of two CG sites. Two important properties of the Hessian arise from eq 10. First, **H** is real and symmetric by construction and is therefore diagonalized by an orthogonal transformation. Where **H** not symmetric, its eigenvectors would not form an orthonormal basis over the full space of molecular motions and NMA could not be performed. Second, none of the eigenvalues of **H** can be negative if **H** is constructed at a local potential energy minimum. The sign of a given eigenvalue indicates the local curvature of the potential along the corresponding mode directional vector or eigenvector: Positive eigenvalues indicate local minima, and negative eigenvalues, local maxima. The local potential energy landscape for a system in a potential energy minimum will have only positive or zero curvature in all directions. Eigenvalues that are identically zero indicate conformational changes that have no effect on the system's (internal) potential energy. Typically, **H** has six zero eigenvalues, corresponding to the rigid-body rotations and translations of the molecule. Obtaining the proper form of the Hessian can be a difficult process that must be handled delicately and will be discussed later.

2.2.3. Equation of Motion and Its Solution—The Hessian does not contain the full story of NMA. Because NMA is applied to the study of dynamics, it is necessary to account for kinetic energy as well as potential energy. In doing so, the form of the matrix that is to be diagonalized changes slightly, but the physical interpretation of the results is more transparent. By considering the system to be a collection of classically behaving particles, the equation of motion can be written as

$$\mathbf{M}\frac{\mathrm{d}^2\Delta\mathbf{q}}{\mathrm{d}t^2} + \mathbf{H}\Delta\mathbf{q} = 0 \quad (11)$$

Here the diagonal matrix **M** contains the masses of the particles. Each mass is repeated three times, once for each of the particle's three Cartesian coordinates. A solution to eq 11 is the 3*N*-dimensional vector $\mathbf{u}_k(t) = \boldsymbol{a}_k \exp\{-i\omega_k t\}$, where $\mathbf{a}_k$ is a complex vector containing both amplitude and phase factor, and ω_k is the frequency of the mode of motion represented by $\mathbf{u}_k(t)$. Substituting this solution into eq 11, the equation of motion becomes

$$\mathbf{H}\mathbf{u}_k = \omega_k^2 \mathbf{M}\mathbf{u}_k \tag{12}$$

which is a generalized eigenvalue equation. The complete set of solutions $\mathbf{u}_k(t)$, $1 \le k \le 3N$, and the corresponding squared frequencies ω_k^2 may be organized as the respective columns of the matrix **U** and the elements $\lambda_k = \omega_k^2$ of the diagonal matrix **Λ** to rewrite the set of 3*N* equations represented by eq 12 in compact notation as

$$\mathbf{H}\mathbf{U} = \mathbf{M}\mathbf{U}\mathbf{\Lambda} \tag{13}$$

Equation 13 is usually solved by transforming it to a standard eigenvalue equation $\tilde{\mathbf{H}}\tilde{\mathbf{U}} = \tilde{\mathbf{U}}\mathbf{\Lambda}$ in mass-weighted coordinates through the transformations $\tilde{\mathbf{U}} = \mathbf{M}^{1/2}\mathbf{U}$ and $\tilde{\mathbf{H}} = \mathbf{M}^{-1/2}\mathbf{H}\mathbf{M}^{-1/2}$. The mass-weighted Hessian, $\tilde{\mathbf{H}}$, retains the symmetry of the original Hessian, and its eigenvectors $\tilde{\mathbf{u}}_k = \mathbf{M}^{1/2}\mathbf{u}_k$ form an orthonormal basis set (i.e., $\tilde{\mathbf{U}}^{\mathrm{T}}\tilde{\mathbf{U}} = \mathbf{1}$). These are the *normal modes* of the system. Their Cartesian counterparts are found through the inverse transformation $\mathbf{U} = \mathbf{M}^{-1/2}\tilde{\mathbf{U}}$ and satisfy the orthonormality condition $\mathbf{U}^{\mathbf{T}}\mathbf{M}\mathbf{U} = 1$. If the interaction sites have all equal mass *m*, **M** reduces to the identity matrix multiplied by *m*, $\tilde{\mathbf{U}} = m^{1/2}\mathbf{U}$, and $\tilde{\mathbf{H}} = m^{-1}\mathbf{H}$.

2.2.4. Significance of Normal Modes and Dominance of Slow Modes—The energy associated with a given normal mode is directly proportional to the square of its frequency (or its eigenvalue $\lambda_k = \omega_k^2$). This can be seen by rewriting eq 9 for a single mode *k*:

$$V(\mathbf{u}_k) = \frac{1}{2}\mathbf{u}_k^{\mathrm{T}}\mathbf{H}\mathbf{u}_k = \frac{\omega_k^2}{2} \tag{14}$$

Displacements along high-frequency modes are therefore energetically more expensive than those of equal magnitude along low-frequency modes. The vibrational energy is, on average, equally partitioned among all the modes, such that the average amplitude of oscillation along mode *k* scales with $1/\omega_k^2$. Thus, the molecule experiences the greatest displacement along the lowest frequency, or "slowest", modes. Conceptually, the energy landscape slopes most gently along the slow modes, and these are consequently the most accessible. These modes are also of highest interest when seeking to determine the most probable *global* fluctuations of a molecule. Large eigenvalues, on the other hand, indicate directions of steep energetic ascent, and excursions along these modes will quickly raise the system's energy.

The cross-correlations $\langle \Delta q_i \Delta q_j \rangle$ between the displacements of the interaction sites along different coordinates are calculated as statistical mechanical averages of the form

$$\left\langle \Delta q_i \Delta q_j \right\rangle = \frac{1}{Z}\int d^{3N} q e^{-1/\left(2k_{\mathrm{B}}T\right)\Delta\mathbf{q}^{\mathrm{T}}\mathbf{H}\Delta\mathbf{q}} \Delta q_i \Delta q_j = k_{\mathrm{B}}T\left(\mathbf{H}^{-1}\right)_{ij} \tag{15}$$

using the configurational integral

$$Z=\int d^{3N}q e^{-1/(2k_{\mathrm{B}}T)\Delta\mathbf{q}^{\mathrm{T}}\mathbf{H}\Delta\mathbf{q}} =(2\pi k_{\mathrm{B}}T)^{3N/2}\left[\det\left(\mathbf{H}^{-1}\right)\right]^{1/2} \quad (16)$$

Here the integrations are performed over the complete space of equilibrium fluctuations, k_B is the Boltzmann constant, T is the absolute temperature, and $(\mathbf{H}^{-1})_{ij}$ designates the ij^{th} element of the inverse of **H**. Because only internal motions affect the system's potential energy, **H** has exactly six eigenvalues that are identically zero, corresponding to the three translational and three rotational degrees of freedom. The inverse of **H** is therefore replaced by the *pseudoinverse*, which is the inverse evaluated only in the space corresponding to the nonzero eigenvalues,

$$\tilde{\mathbf{H}}^{-1}=\sum_{k=1}^{3N-6}\frac{\tilde{\mathbf{u}}_k\tilde{\mathbf{u}}_k^{\mathbf{T}}}{\omega_k^2} \quad (17)$$

The importance of the slow modes is again highlighted in these equations: The lowest frequency modes contribute most to the spatial partition function because $\det(\tilde{\mathbf{H}}^{-1})$ is the product of the reciprocal nonzero eigenvalues of $\tilde{\mathbf{H}}$.

2.2.5. Covariance Computed from NMA: Bridging with PCA of Structural Ensembles—The cross-correlation $\langle\Delta q_i\Delta q_j\rangle$ on the left-hand side of eq 15 is simply the ij^{th} element of the covariance matrix **C**; therefore, eq 15 may be rewritten in compact notation as

$$\mathbf{C}=k_{\mathrm{B}}T\mathbf{H}^{-1}$$

This equation establishes the bridge between the PCA of ensembles of conformations and NMA of a given structure. In the former case, the top ranking (principal) modes of structural changes are extracted from experimental data (or sets of known structures for a given protein). In the latter, the same such structural changes are *predicted* by the theory using one structure to construct **H**.

The top-ranking modes obtained by PCA should, in principle, be comparable to the lowest frequency modes derived by NMA (i.e., $\lambda_i \sim 1/\sigma_i$, and $\mathbf{p}_i \sim \mathbf{u}_i$), provided that (i) the data set of conformations subjected to PCA represents an equilibrium distribution and (ii) the Hessian in NMA provides an accurate description of dominant interactions. Recent PCAs performed for ensembles of PDB structures exhibit good agreement with the global modes predicted by CG NMAs.[44,190] Notably, ENMs have been adopted in those NMAs. The relevance of ENM predictions for a given protein to PC modes derived from sets of structures experimentally resolved for the same protein (under different conditions, in the presence of different ligands) lends support to the use of ENMs for assessing functional changes in structure. Similar results will be presented below for rhodopsin.

2.2.6. Using Normal Modes for Exploring the Potential Energy Surface—The harmonic approximation only holds in the immediate vicinity of a local potential energy minimum, but what if we wish to explore structures that are far away from this minimum? One method for exploring remote regions of the potential energy surface is the normal mode following (NMF) technique.[191] In this method, one starts at an energy minimum and iteratively

moves the structure along its slowest eigenmode while remaining at a minimum for all the other modes. Eventually one of the eigenvalues will become negative, indicating the neighborhood of a saddle point or a transition state. From that point, other local minima can be found by iteratively distorting the structure along the potential energy gradient.

The NMF method has recently been enhanced through using the Metropolis MC criterion to control the size of the steps taken.[192] As discussed in section 3.1.1, this technique has proven useful in revealing the gating mechanism of gramicidin A. Similar techniques that take advantage of movements along the mode coordinates have been exploited for investigating transition pathways between known minima, as described in section 4.4.

2.3. Elastic Network Models

NMA requires knowledge of a symmetric and nonnegative Hessian. An energy minimization is required prior to performing NMA on a protein crystal structure to ensure that the first derivative of the total potential is zero with respect to all degrees of freedom and to evaluate the second derivatives (elements of **H**). Energy minimization is a computationally expensive task and generally distorts the initial conformation, resulting in NMA being performed on a structure altered from the original. Lu and Ma have demonstrated that the problem of initial energy minimization can be overcome by mathematically moving the minimum to the initial structure.[193] Their technique involves decomposing the Hessian into submatrices, replacing each submatrix with its nearest symmetric positive semidefinite matrix, and reconstructing the Hessian. Far easier is adopting an ENM that by design accepts the initial structure, usually taken from the PDB, to be an energy minimum.

The ENM representation is readily scalable to any level of coarse-graining and requires very few parameters. The ENM approximates the protein's potential energy as that of a classical network of masses coupled by springs: each node in the network is a CG site, and each edge is a spring. The network topology is defined by the native structure, with edges placed between nodes/sites that lie within a prespecified cutoff distance from each other. Comparisons of predicted rms fluctuations to motions inferred from crystallographic B-factors have identified optimal cutoff distances of 7.3 Å for the GNM and 18 Å for the ANM, provided that nodes are identified by the C^{α}-atoms.[78,194] As to the spring constants, the simplest ENMs use a uniform force constant for all interactions; Hinsen proposed using a force constant that decays rapidly with distance.[74] Sen and Jernigan empirically investigated how the force constants should vary with the residues' coordination numbers.[195] The adoption of stiffer springs for sequentially neighboring residues[196] or amino acid-specific force constants[197,198] has been shown to improve the agreement with experiments.

The choice of the specific spring constants has little, if any, effect on the global modes. The global modes of motion are widely recognized to be intrinsic properties of the 3D shape of the protein and have been verified in several studies to be insensitive to model parameters[61–66] and almost identically reproduced at various hierarchical levels of resolutions.[45,61,199] The robustness of global modes permits us to utilize ENMs in the study of membrane proteins. One might conceivably adopt different force constants for the internal and interfacial regions of membrane proteins and even differentiate between the interactions with lipid molecules and those with water molecules in the EC or CP regions. However, as will also be illustrated below, comparisons of ENM results with those obtained from full atomic NMAs conducted in the presence of explicit water/lipid molecules have shown that the global modes of membrane proteins are essentially dictated by the protein architecture/fold/shape, similar to the cases for other proteins, and are robust to small variations in the EN topology and environmental effects. Furthermore, the structural changes along the global modes are observed to correlate well with those experimentally observed for particular membrane proteins that are structurally characterized in different states (e.g., apo vs ligand-bound forms).

Several ENM servers have been developed to date, which permit users to readily retrieve results based on the ENMs and their extensions to several applications.[78,186,200–205] Below we present the theory and assumptions underlying commonly used ENMs, and sections 3 and 4 will illustrate their applications and extensions.

2.3.1. Gaussian Network Model—The GNM is based on the assumption that all residue fluctuations (and inter-residue distances) are Gaussianly distributed around their equilibrium coordinates,[67,68] similar to the statistical mechanical behavior of polymer networks.[69–73] The equilibrium coordinates are identified by the position vectors $\mathbf{R}_i^0$ of C^α-atoms in the PDB structures. Residue pairs are connected by a spring of force constant γ_{ij}, provided that they are located within a cutoff distance R_c. The fluctuations in residue positions and their cross-correlations are fully controlled by the $N \times N$ Kirchhoff matrix, $\mathbf{\Gamma}$, defined in terms of the spring constants as

$$\Gamma_{ij} = -\gamma_{ij} \tag{19}$$

for all $i \neq j$, and

$$\Gamma_{ii} = \sum_j \gamma_{ij}$$

where the summation is performed over all off diagonal terms in the row i (or column j). $\mathbf{\Gamma}$ is the $N \times N$ counterpart of $\mathbf{H}$. Its pseudoinverse, $\mathbf{\Gamma}^{-1}$, scales with the covariance matrix $\boldsymbol{C}$[67,68,206] (see eq 5)

$$\bar{\mathbf{C}} = 3k_B T \Gamma^{-1} \tag{20}$$

The above equation is obtained from a statistical mechanical average similar to eqs 15 and 16, where the overall interaction potential is replaced by[207,208]

$$\begin{aligned} V_{GNM} &= \frac{1}{2}\sum_{ij}\gamma_{ij}\left(\mathbf{R}_{ij} - \mathbf{R}_{ij}^0\right)\left(\mathbf{R}_{ij} - \mathbf{R}_{ij}^0\right) \\ &= \frac{1}{2}\sum_{ij}\gamma_{ij}\left(\Delta\mathbf{R}_{ij}\right)^2 \end{aligned} \tag{21}$$

The above summation is performed over all connected pairs. A major simplification in the GNM is the adoption of a uniform spring constant $\gamma_{ij} = \gamma$ for all residue pairs (i, j) that are connected. As a result, $\mathbf{\Gamma}$ reduces to the adjacency matrix, or Lagrangian, multiplied by γ. Note that the absolute value of γ does not affect the mode shapes (or eigenvectors) but uniformly scales their squared frequencies (eigenvalues).

To date, the GNM has been tested in numerous applications and proven to yield results in reasonable agreement with a wealth of experimental data, including X-ray crystallographic *B*- factors for amino acids[194,204] and nucleotides,[205] root-mean-square deviations in residue coordinates for NMR models,[209] H/D exchange free energy costs,[210] hinge sites in many enzymes and their spatial proximity to catalytic sites,[32] NMR order parameters[211,212] and changes in NMR parameters upon ligation,213 highly conserved core amino acids,206 unfolding pathways214 and folding nuclei[215] in proteins (e.g., rhodopsin),[216] or the common

dynamics of families of proteins applied to globins,217 and potassium channels.218 The good correlation between GNM predictions and experimental data observed in numerous applications despite the simplicity of the model highlights the important role of native contact topology in defining the collective dynamics.

The eigenvalue decomposition of **Γ** permits us to assess the contribution of different modes to equilibrium dynamics. **Γ** has $N - 1$ nonzero eigenvalues, with the lowest corresponding to the first (global) mode. Typical outputs from GNM mode decomposition include the displacement of residues along each mode axis (global hinge sites being located between sequence segments that undergo opposite direction movements along slowest modes), cross-correlations between residues in individual modes, and square displacement profiles of residues driven by individual modes or subsets of modes. No information on the 3-dimensional directions of motions can be obtained with the GNM, because the main ingredient of the theory is an $N \times N$ matrix (as opposed to the $3N \times 3N$ Hessian in NMA). The anisotropic network model, described next, is the simplest ENM that provides information on directionalities.

2.3.2. Anisotropic Network Model—The most broadly used ENM is the anisotropic network model.[76–79] The positions of the nodes in the ANM are identified by the coordinates of C^{α}-atoms for amino acids, and P, C_{4*}, and C_2-atoms for nucleotides. ANM analysis is simply a CG NMA, subject to the potential[74,75]

$$V=\frac{1}{2}\sum_{ij}\gamma_{ij}\left(R_{ij}-R_{ij}^{0}\right)^{2} \tag{22}$$

Note that there is a central difference between V_{GNM} and the above potential. Here $V = 0$ if $R_{ij} = R_{ij}^0$, irrespective of the direction of the corresponding distance vectors. In the case of the GNM, on the other hand, changes in the distance *vector* incur a potential energy increase, even if the inter-residue distance is maintained (see eq 21). The mean-square fluctuations and cross-correlations predicted by GNM have been shown in comparative studies to yield better agreement with experimental data than the ANM predictions.[38,39,207,208]

Using the ANM, it is possible to readily write a closed form expression for **H** using eq 22 in eq 10. The second derivatives of the potential in this case are simply given by

$$\frac{\partial^{2}V}{\partial x_i\ \partial y_j}=-\frac{\gamma_{ij}\left(x_j-x_i\right)\left(y_j-y_i\right)}{R_{ij}^{2}} \tag{23}$$

Using the notation $x_{ij}^0 = (x_j^0 - x_i^0)$ and similarly for y_{ij}^0 and z_{ij} for the three components of the instantaneous distance vector $\mathbf{R}_{ij}^0$, the off-diagonal 3×3 submatrices of **H** take in the ANM the form

$$\mathbf{H}_{ij}=-\frac{\gamma_{ij}}{\left(R_{ij}^{0}\right)^{2}}\begin{bmatrix}\left(x_{ij}^{0}\right)^{2} & x_{ij}^{0}y_{ij}^{0} & x_{ij}^{0}z_{ij}^{0}\\ x_{ij}^{0}y_{ij}^{0} & \left(y_{ij}^{0}\right)^{2} & y_{ij}^{0}z_{ij}^{0}\\ x_{ij}^{0}z_{ij}^{0} & y_{ij}^{0}z_{ij}^{0} & \left(z_{ij}^{0}\right)^{2}\end{bmatrix} \tag{24}$$

and the diagonal submatrices satisfy the identity

$$\mathbf{H}_{ii} = -\sum_{j;j\neq i} \mathbf{H}_{ij} \tag{25}$$

This simple expression for **H** is readily used in NMA to determine the collective dynamics. We note that the amino acid specificity can be included in ENM-based studies by adopting residue-specific force constants, and indeed we have deliberately presented the GNM Kirchhoff matrix and ANM Hessian (respective eqs 19 and 24) in terms of force constants, γ_{ij}, that are dependent on the identity of the amino acids *i* and *j* connected in the network. However, in most applications, γ_{ij} is taken as a constant, γ, for all pairs of residues connected in the network. Equation 22 with a single parameter $\gamma_{ij} = \gamma$ has been originally used by Tirion for representing *interatomic* interactions (as opposed to *inter-residue* interactions considered in all succeeding ENM studies, starting from the GNM) and demonstrating the reproducibility of global modes obtained by detailed atomic force fields.[58] As mentioned above, the absolute value of γ for a given level representation does not affect the mode shapes (i.e., the eigenvectors, $\boldsymbol{u_k}$, $(1 \leq k \leq 3N-6)$ of **H**) but their frequencies, because the eigenvalues of **H**, λ_k, are proportional to γ. Likewise, the global modes are insensitive to the adoption of residue-specific force constants. A more detailed assessment of the specific role of particular residues in these global modes and the redistribution of interactions (e.g., salt bridges) resulting from global movements, and their effect on allosteric pathways will be given below.

A major utility of the ANM is its ability to generate alternative conformations (substates or microstates) in the close neighborhood of a given structure upon deforming the original structures along the dominant (lowest frequency) modes. Similar to eq 7, the change in the square fluctuations of residue *i* contributed by the movement along a given mode *k* is given in terms of the *k*th eigenvector ($\mathbf{u}_k$) and eigenvalue (λ_k) of **H** as

$$(\Delta \mathbf{R}_i)^2\Big|_k = \mathrm{tr}\left\{\left[\lambda_k^{-1}\mathbf{u}_k\mathbf{u}_k^{\mathrm{T}}\right]_{ii}\right\} \tag{26}$$

Or the alternative conformations induced upon moving along a given mode are simply[31]

$$\mathbf{q}^{(k)} = \left(\left(\mathbf{R}_1^{(k)}\right)^{\mathrm{T}}, \ldots, \left(\mathbf{R}_N^{(k)}\right)^{\mathrm{T}}\right)^{\mathbf{T}} = \left(\left(\mathbf{R}_1^{0}\right)^{\mathrm{T}}, \ldots, \left(\mathbf{R}_N^{0}\right)^{\mathrm{T}}\right)^{\mathbf{T}} \pm \mathbf{s}\lambda_k^{-1/2}\mathbf{u}_{\boldsymbol{k}} \tag{27}$$

where the coefficient *s* scales with $(k_B T)^{1/2}$. In principle, given the uncertainty in the absolute value of γ, which is reflected on the eigenvalues, a range of *s* values giving rise to movements comparable in size to those experimentally observed may be generated and used for further calculations (such as generating an ensemble of conformations to be used in docking simulations; see section 4.2). Alternatively, the choice of *s* may be based on the correlation cosine or *overlap*[219]

$$I_k = (\Delta \mathbf{q}_{\mathrm{AB}} \bullet \mathbf{u}_k) / |\Delta \mathbf{q}_{\mathrm{AB}}| \tag{28}$$

between the normalized directional vector $\mathbf{u}_k$ and the targeted direction of deformation $\Delta\mathbf{q}_{\mathrm{AB}} = \mathbf{q}^{(\mathrm{B})} - \mathbf{q}^{(\mathrm{A})}$, provided that the goal is to explore the transition from substate A to B. The potential contribution of subsets of modes to such a transition may be deduced from the *cumulative overlap* $[\sum_k I_k^2]^{1/2}$, where the summation is performed over the subset of modes of

interest, usually starting from the lowest-lying modes. Note that this summation is identically equal to unity if it is performed over all $3N - 6$ modes/eigenvectors, which form a complete orthonormal basis set for the $3N - 6$ dimensional space of conformational changes. Another quantity of interest is the *degree of collectivity*, κ_k, for mode k, defined as[220]

$$\kappa_k = \boldsymbol{N}^{-1} \exp\left\{ -\sum_{i=1}^{N} \alpha \, (\Delta \mathbf{R}_i)^2 \Big|_k \log(\alpha \Delta \mathbf{R}_i)^2 \Big|_k \right\} \tag{29}$$

where α is the normalization constant $\sum_i \alpha(\Delta \mathbf{R}_i)^2|_k = 1$. The form of eq 29 suggests that the degree of collectivity has an entropic significance. The mode with the highest degree of collectivity has the highest entropy: it is distributed over a larger number of residues rather than being orderly confined to a few residues. Lower frequency modes are usually more collective; their high degree of collectivity is indeed needed for triggering cooperative (allosteric) responses. Of interest is to identify the most collective modes toward disclosing potentially functional movements intrinsically favored by the overall structure. Sections 3.1, 3.2, and 3.3 will present applications of the ANM to ion channels, receptors, and transporters, respectively.

2.3.3. Rotating-Translating Blocks Model—A key strength of ENMs is their scalability. Because the interactions are all pairwise and harmonic, once the CG sites are defined, the ENM can be constructed and its Hessian determined. Scalability is particularly useful when modeling very large systems, as it is often the case that the memory required for diagonalizing **H** exceeds that currently accessible. The slow modes predicted by NMA are robust to the level of coarse-graining, and bundling 20 or more residues into a single CG site still produces slow modes that overlap well with the global modes of the full C^α representation.[61,199] The disadvantage of excessive coarse-graining is the loss of information on the detailed local movements. Although global motions are accurately reproduced with high levels of coarse-graining, reconstructing their details can be daunting. Mixed models[221–223] that use detailed descriptions only for specific regions of interest and CG descriptions for most of the molecule are helpful in retaining desired detail while discarding unnecessary information.

In special cases, the size of the Hessian can be reduced by exploiting the symmetry of the system. Group theoretical calculations were used to represent the Hessian of icosahedral viral capsids in reduced forms.[224–226] Alternatively, by making the assumption that all repeat units in a symmetric system behave identically, one can construct a reduced Hessian that has only symmetrical modes.[227] A more general method for reducing the complexity of **H** without eliminating any structural detail is the rotations and translations of blocks (RTB)228 or the block normal mode229 (BNM) method. This method assumes that the system is constructed of n_b rigid blocks and that the normal modes can be expressed as rigid body rotations and translations of its constituent blocks. Each block has six degrees of freedom (three translational, three rotational). The number of degrees of freedom thus reduces from $3N$ to $6n_b$. The blocks are defined as seen fit for the application at hand: An all-atom protein model might be simplified by assuming that each residue forms a rigid block, or a C^α-only model might be simplified into blocks of secondary structure. Furthermore, the size of the blocks is not restricted: If some domain is known to be particularly rigid, it might be modeled as a block, whereas a small but flexible loop may consist of several blocks. The limitation of the RTB method is that it does not reproduce internal motions of the blocks, so that a great deal of information can be lost if flexible regions with high internal mobility are assumed to be rigid.

Consider a system of N particles that can be collected into $n_b < N$ rigid blocks connected by elastic springs. Define the $3N \times 6n_b$ projection matrix, **P**, from the $3N$-dimensional space of

all particles into the $6n_b$-dimensional space of rotations and translations of the rigid blocks. The original Hessian is projected into the space of rigid blocks with the transformation

$$\mathbf{H}_{\mathrm{BLK}} = \mathbf{P}^{\mathrm{T}}\mathbf{H}\mathbf{P} \tag{30}$$

$\mathbf{H}_{\mathrm{BLK}}$ is diagonalized with $\mathbf{V}_{\mathrm{BLK}}{}^{\mathrm{T}}\mathbf{H}\mathbf{V}_{\mathrm{BLK}} = \mathbf{\Lambda}_{\mathrm{BLK}}$, and the resulting eigenvectors are projected back into the full $3N$-dimensional space with the inverse projection $\mathbf{V} = \mathbf{P}^{\mathrm{T}}\mathbf{V}_{\mathrm{BLK}}$. Thus, $6n_b - 6$ normal modes result from the rigid block approximation. Each mode is $3N$ dimensional.

This method was first applied to small proteins by Durand et al.,[230] who used it to simplify conventional MD force fields by grouping atoms into rigid amino acids. It has since been used to investigate the role of intrinsic dynamics in conformational changes in molecular motors,[229,231] to study the motion of the ribosome,[81] the maturation of cowpea chlorotic mottle virus,[232] and the mechanical properties of icosahedral viral capsids.[233]

2.3.4. Extensions for Treating Environmental Effects—Methods based on ENMs have been enhanced to include the effect of viscous drag,[234–236] and altered to include the environmental perturbations.[28,237] In order to systematically assess the effect of environment on protein dynamics, Ming and Wall[28,238] and Zheng and Brooks[237] proposed a method that relies on separating the problem into a *system* that contains all relevant degrees of freedom and an *environment* that contains all other degrees of freedom. The Hessian matrix is then composed of four blocks that relate the system with itself ($\mathbf{H}_{ss}$), the environment with itself ($\mathbf{H}_{ee}$), and the system with the environment ($\mathbf{H}_{se}$),

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{ss} & \mathbf{H}_{se} \\ \mathbf{H}_{se}^{T} & \mathbf{H}_{ee} \end{pmatrix} \tag{31}$$

At a minimum of the potential energy, the pseudo-Hessian, $\bar{\mathbf{H}}$, is found as

$$\bar{\mathbf{H}} = \mathbf{H}_{ss} - \mathbf{H}_{se}\mathbf{H}_{ee}^{-1}\mathbf{H}_{se}^{T} \tag{32}$$

$\bar{\mathbf{H}}$ has the same dimension as $\mathbf{H}_{ss}$ but includes the effects of the environment. Its eigenmodes can be directly compared to those of any system of equal size. This technique has been used to study a range of phenomena, including the coupling of motor protein binding pocket dynamics to global protein structure,[237] substrate induced conformational changes,[239] and allostery in membrane proteins.[29,240]

Another method for introducing viscous damping into a vibrational system is to use the Langevin equation,

$$\mathbf{M}\frac{\mathrm{d}^2\Delta\mathbf{q}}{\mathrm{d}t^2} + \mathbf{Z}\frac{\mathrm{d}\Delta\mathbf{q}}{\mathrm{d}t} + \mathbf{H}\Delta\mathbf{q} + \xi(t) = 0 \tag{33}$$

Here the elements of the friction matrix, $\mathbf{Z}$, provide velocity-dependent damping, and the white noise vector ξ accounts for thermal energy transferred to the molecule from the solvent. The elements of this vector obey the properties

$$\langle \xi_i(t) \rangle = 0 \tag{34}$$

$$\left\langle \xi_i(t) \quad \xi_j(t') \right\rangle = 2Z_{ij}k_B T \quad \delta(t-t') \tag{35}$$

From eq 34 it is seen that the net external force incident on each CG center averages to zero. Equation 35 indicates that the external force is random in time and provides as much energy as is lost due to damping. The solution to eq 33 for macromolecules was given by Lamm and Szabo[241] and has further been modified to incorporate the use of rigid blocks.[83] When compared with MD, the Langevin models provide insight into the role of friction in protein dynamics.[234,236,242] This technique has been used in conjunction with ENMs to calculate scattering functions of proteins,[243] to investigate the sources behind damping in global protein motions,[244] and to estimate the fractional free energy loss in the myosin power stroke.[235]

We note that, in a related study,[83] the response of membrane-embedded gramicidin A dimer to a sudden velocity kick near one end was explored by examining the time evolution of the molecule, modeled as a collection of harmonic oscillators, under the Langevin equation. Calculations shortly referred to as Langevin dynamics (conceptually similar to NMA, but in the phase space of displacements and momenta) were repeated twice, for the fully atomic model and for a mixed model where the eight indole groups on the molecule were represented by the RTB model. Detailed comparison of the results from the two sets of calculations showed that the results from RTB-Langevin dynamics closely agreed with those from full atomic Langevin dynamics, in support of the adoption of the RTB model for structural elements that are known to be nearly rigid.

3. Intrinsic Dynamics of Membrane Proteins and Their Functional Significance

3.1. Ion Channels

Ion channels are usually accepted to be passive transport proteins: they allow for the conduction of ions when the electrochemical gradients are shifted away from the equilibrium membrane potential. Their functions include establishing a resting membrane potential, controlling cell volume, and regulating the flow of ions across the epithelial cell.[89] Their role of “facilitator” is achieved by undergoing changes between open and closed conformations. Fundamental questions associated with ion channel functions concern the location of the activation gate as well as the conformational changes that ensure the reversible occlusion of the channel. Here, we will present the NMA results for gramicidin A, potassium channels, and MscL.

We note that, among the global modes predicted for symmetric structures (e.g., homotetrameric potassium channels and homopentameric MscL), a group of modes maintain the structural symmetry; that is, they induce the same type of deformation in all monomers. These modes are nondegenerate; that is, they have unique eigenvalues. Nondegenerate modes play a dominant role in enabling the cooperative transitions of large multimeric, structurally symmetric proteins (or assemblies), by simultaneously exploiting the intrinsic preferences of individual subunits. It will be shown below that such nondegenerate modes are instrumental in initiating the gating process in ion channels.

3.1.1. Gramicidin A—Gramicidin A is the first membrane protein that has been examined by NMA.[245] It is also the smallest ion channel known to date: it is a dimer, with each monomer being simply a left-handed helix of 16 amino acids. The two helices are stacked head-to-head and allow for the selective permeation of small cations (e.g., Cs^+, K^+) and water molecules through a narrow opening along the helical axes (Figure 7A). The original NMA of GA dimer

by Roux and Karplus was one of the earliest studies of membrane proteins, if not the first, that overruled the use of rigid, or nearly rigid, models as appropriate approximations to simulate ion permeation.[245] The GA was demonstrated therein to enjoy considerable structural flexibility. Another interesting observation, which probably was one of the early indications of the robustness of low frequency modes, was the observed insensitivity of the low frequency modes to changes in the strength of hydrogen bond interactions. Computations performed by varying the partial charges assigned to carbonyl dipoles showed that the frequencies below 75 cm^{-1} remained virtually unchanged; and since these low frequency modes had a dominant effect on the fluctuation behavior, the overall GA dynamics exhibited little dependence on the strength of hydrogen bond interactions.

Extensive computational studies have been performed for gramicidin since then.20 The gating mechanism and the slow conformational transitions undergone by GA have now been identified to be a counter-rotation of the two helices around the pore axis, accompanied by a slight expansion of the channel mouths at the EC and CP ends. To elucidate this mechanism, Miloshevsky and Jordan performed192 a series of computations, including MC-NMF (see section 2.2.6) coupled with a Metropolis algorithm, full atomic NMA, NMA with the one-residue-per-block RTB approach, and NMA with the ANM (using the elNémo server202). Notably, the global mode predicted by these models of various complexities was invariably observed to be the counter-rotation of the two helices around the pore axis (Figure 9). The global mode shape obtained with the ANM was reported therein to look “as accurate as that obtained with the all-atom CHARMM22 force field”. It is worth noting, however, that the global mode frequency predicted by the RTB and ANM is significantly larger than that (~6.5 cm^{-1}) observed in full atomic NMA. Therefore, while the functionally relevant *mechanism* of motion can be accurately predicted by these CG models, the time scales are *not*. This is understandable, as the slow modes are dampened by solvent and these CG models do not take account of the viscous drag effect.

The conservation of the mode shape between ANM- and CHARMM-based NMAs is in accord with compelling evidence that the directionality of the structural changes natively accessible to proteins is not altered by the solvent effect or the use of highly simplified potential functions.[66] Miloshevsky and Jordan concluded that the predicted gating mechanism is an inherent property of GA architecture and should not be changed by the surrounding lipid and water molecules.[192]

It is also worth noting that in contrast to the generally accepted gating model of two states, closed and open, involved in GA's functional transitions, the simulations show that the GA may exist in multiple intermediate states, consistent with experiments.[246,247] Finally, the effective sampling of the transition state using the MC-NMF supports the view that the movements along the global mode predicted by NMA (either full atomic or using ANM) form the crucial step for initiating pore opening. Motions along the lowest eigenvalue modes encountered at the later stages of the transition pathway via adaptive NMF exhibited some departures, however, from those predicted by RTB and ANM for the initial state.

3.1.2. Potassium Channels—Potassium channels are tetramers, cylindrically arranged to form a bundle of TM helices, enclosing a central pore, or a channel, through which ions are conducted. The pore regions of most K^+ channels are considered to have similar structure, despite significant differences in sequence (Figure 10). They all contain two TM helices, TM1 (yellow) and TM2 (blue), per monomer, connected by a stretch of 30 residues, known as the P-loop region (red). The P-loop contains three structural elements: a narrow selectivity filter of ~10 Å length near the EC entrance of the pore region; the P-helix, which spans only the upper half of the bilayer; and the exposed loops, also known as the turret, at the EC side. The selectivity filter is followed by a large cavity in the middle of the core region, which ends in a

CP gating region, as illustrated in Figure 11A for KcsA, the first K^+ channel that has been crystallized and structurally resolved.[118] The outer helices (TM1) are exposed to the lipid environment; the inner helices (TM2) line the pore. The four P-loops together form the EC vestibule, which opens up into a large central aqueous cavity (of ~10 Å diameter in KcsA).

The KcsA structure is proposed to be in the closed form.[156] The structure of a calcium-dependent K^+ channel from *Methanobacterium thermoautrophicum* (MthK)[120] is considered to be the open form, containing a wide open intracellular pore of ~16 Å. Since the determination of the KcsA structure,118 many more K^+ channel structures have been resolved in either open or closed forms (Figure 10).

The availability of these structures allows NMA-based studies exploring the collective movements of the potassium channels and assessing, in particular, the pore opening mechanism. One of the first studies in that direction was a NMA of KcsA by Ma and co-workers, in 2002.[249] The study pointed for the first time to the concerted rotational motion of all four TM2 helices as a collective mode favored by the tetrameric structure. More recent examination of a series of potassium channels using the ANM demonstrated that the core domains favor exactly the same mechanism of global motion in all cases, which allows for pore opening.[218] This global mode of motion is a counter-rotation of the two halves of the molecule about the cylindrical axis of symmetry, akin to a concerted *twisting-and-torsion* motion of all TM helices. This nondegenerate mode equally distorts all four subunits and confers a remarkable expansion at the gate region (Figure 11B) by swinging the M2 helices away from the cylindrical axis, while the selectivity filter region remains fairly rigid. A striking observation is the appearance of a kink in the TM2 helices which further enhances the pore opening.[218]

Notably, the movement of the M2 helices is consistent with the displacement observed in the MthK crystal structure.120 Furthermore, the change in the relative positions of the four TM2 helices at the gate is also consistent with the spin-labeling experiments of Perozo et al.157, 158 and single molecule techniques[250] which point to an increase in the distances between the M2 helices (or diameter) at the pore region and a kink at G83 in MthK TM2 (counterpart of G99 in KcsA and G134 in KirBac1.1; see Figure 10A), in accord with ANM results. The comparison of the pore-radius profiles for the wild type protein and its “deformed” form predicted by the ANM in Figure 11B clearly illustrates the increase in the pore-radius at the CP gate region.

The recently resolved X-ray structure of a nonselective cation channel NaK (PDB ID: 3E86) in the open form[251] provides an elegant example of the role of kink-formation in opening up the cation channel. This structure, when superimposed onto the closed form (PDB ID: 2AHY) 252 indicates the selectivity filter to be static during gating,251 in agreement with ANM predictions218 and SDSL-EPR measurements157,[158] described above. The major conformational change is a helix bending at the highly conserved G87, which acts as a hinge. The counterpart of this glycine, conserved in potassium channels’ TM2 helices (Figure 10), has been pointed out to act as a hinge site,[218] which also exhibits a kink during the gating motion. Furthermore, a comparison of the open and closed forms of NaK also shows a global-twisting motion around the helical axis of the inner helix,[251] in agreement with experiments[250] as well as ANM predictions.[218]

Miloshevski and Jordan applied their MC-NMF method (successfully used in their earlier examination of GA channel gating; see Figure 9) to KcsA. Their study also confirmed that the gating mechanism of KcsA involves a rotation and unwinding of the TM2 bundle away from the channel axis, leading to an open state with an inner vestibule of ~5–7Å radius, in agreement with the computational models described above.189 Haliloglu and Ben-Tal[253] also analyzed

the transition between the closed and open structures (KcsA and MthK, respectively) using the ANM and *in silico* alanine-scanning mutagenesis data. Their ANM study again confirmed the global torsion motion as the dominant mechanism of pore opening, while the alanine-scanning mutagenesis study identified a network of energetically and dynamically coupled residues between the selectivity filter and the CP region, consistent with experimental data.[254]

3.1.3. Mechanosensitive Channels—One of the most basic demands of primitive cells is to tolerate changes in the environment, such as tonicity, without bursting. This function involves regulation of cell volume by ion flow.[255] MscL is an ion channel that is able to detect and relieve such tensions in the membrane.[256–259]

The crystal structure of the homopentameric MscL from *Myobacterium tuberculosis* has been resolved at 3.5 Å by Chang and co-workers.[117] Each subunit contains three α-helices: two TM (TM1 and TM2) and one CP (Figure 12A). An aqueous cavity opening, approximately 18 Å in diameter, leads from the EC side, through a pore lined with hydrophilic residues narrowing down at the CP side, to an occluded hydrophobic apex, which is proposed to be the gate.[117] The five subunits are organized into two domains, the TM and CP domains, both exhibiting a 5-fold symmetry. Their respective diameters are 50 and 15 Å. The pore-lining helix of each subunit, TM1 (yellow in Figure 12A), is connected to the outer helix, TM2 (blue), by an EC loop of 44–68 residues, forming a flap at the EC surface. TM2 is connected to the CP helix by a shorter loop of 10–12 residues. The TM1 and TM2 helical axes are tilted by about 28° with respect to the 5-fold axis while the CP helix is tilted by 15°. The radius of the pore varies between 2 and 18 Å and is partially occluded at the CP region. In the open form, the cross-sectional diameter of the TM domain is estimated to be between 30 and 40 Å,[260] suggesting a substantial increase in the size of the channel pore.

The dynamical properties of MscL from *E. coli* have been explored by NMA[261] using the homology models for the closed form based on the *M. tuberculosis* MscL structure[258,261] and other structural models proposed by Sukharev and co-workers.[258] The study identified two major kinds of motions: type I, a symmetrical motion that corresponds to an overall twisting and tilting of all TM helices around the cylindrical axis, exhibited by the first nondegenerate ANM mode, and type II, a global bending, via modes 2 and 3. The "*twist to open*" motions[261] are consistent with the iris-like mechanism proposed by Sukharev and co-workers[258] to be implicated in the gating process.

Figure 12 displays the structural changes driven by the lowest nondegenerate ANM mode (mode 1), calculated for the *M. tuberculosis* MscL. As indicated by the arrows, the TM domain undergoes a global twist in this mode, with the EC and CP regions undergoing counter-rotations about the cylindrical axis. The amplitude of the motion is higher at the CP ends than at the EC ends. Calculations performed with and without the CP domain show that the same mechanism of motion is maintained at the TM domain, irrespective of the presence or the excision of the CP domain, except for shifting the location of the hinge region by no more than three or four residues.

A small subset of low frequency modes accessible to the initial substate have been shown in previous work to account for 65% of the conformational change observed between the closed and open states.[261] The contribution of the individual modes to the transition between two states may be assessed by examining the overlap (eq 28) between the eigenvectors predicted for a given starting conformation and the difference vector $\Delta \mathbf{q}_{AB} = \mathbf{q}^{(B)} - \mathbf{q}^{(A)}$ between the two end points. Here we focus on the difference vector $\Delta \mathbf{q}_{OC}$ between an open form (O) modeled by Sukharev et al.[258] and the closed form (C). Figure 13 displays the cumulative overlap between $\Delta \mathbf{q}_{OC}$ and the modes predicted by the ANM. ANM calculations were repeated with the open (red dashed curve) and the closed (blue, solid curve) structures. It is clearly seen

that although the pentameric structure has access to $3N - 6$ modes (i.e., 1539 modes in the present case, with $N = 515$), only a small number of modes make a distinctive contribution (see the jumps in the curves). The fact that the transition is achieved by moving in such a small subspace is remarkable. The modes that make the largest contributions are the nondegenerate slow modes. These modes maintain the pentameric symmetry of the channel. In panel B of Figure 13, the motions induced in the second lowest nondegenerate mode (mode 6) are illustrated by the color-coded ribbon diagrams (red, most mobile; blue, most rigid) and arrows along the deformation directions. It can be clearly seen that this motion tends to contract/expand the channel along the cylindrical axis, with the strongest effect exerted on the partially disordered segments exposed to the EC region.

3.2. Receptors

3.2.1. Nicotinic Acetylcholine Receptor—Communication between nerve cells takes place at junctions called synapses. The presynaptic cells release, upon activation, neurotransmitters into the synapse, which bind to ligand-gated ion channels (LGICs) on the surface of the postsynaptic cells. Binding of neurotransmitter causes the channels to open, allowing the ions to flow across the postsynaptic cell membrane. The opening and closing of LGICs rapidly convert chemical signals into an electrical output, regulating the flow of information. Mutations in LGICs lead to a number of "channelopathies", such as congenital myasthenic syndromes, epileptic disorders, and hereditary hyperekplexia.[262] Approximately 8.3% of small-molecule drugs target LGICs (Figure 4).

The nicotinic acetylcholine receptor is a member of a superfamily of pentameric transmitter-gated ion channels, also called Cys-loop receptors, which include the serotonin 5-HT3, GABA A and GABA C, and glycine receptors. Members of this superfamily contain a signature loop of 13 residues closed by a disulfide bridge, called the Cys-loop, at the interface between the EC and TM domains of their respective monomers.[263] The nAChR activity is triggered by binding of acetylcholine (ACh) or nicotine.

The structure of nAChR in the closed state has been determined by cryo-EM of tubular crystals grown from the electric organ of *Torpedo marmorata*.[264,265] The structure consists of five subunits (α, β, δ, α, and γ), two of which (α-subunits) have a slightly distorted (closed or tense, T) conformation compared to the other three (open, relaxed, R), hence the pseudosymmetric organization of the quaternary structure. The receptor is organized into three domains (Figure 7D): a large N-terminal EC domain involved in binding the neurotransmitter, a TM pore domain, and a smaller CP domain. The N-terminal domain of each subunit is composed of an N-terminal α-helix and two β-sheets arranged in a curled β-sandwich connected by the Cys-loop (Figure 14A). The same fold is exhibited by the soluble ACh binding protein (AChBP).[266] There are two ACh binding sites at the interfaces between the α–δ and α–γ subunits' EC domains. The TM domains of individual subunits are composed of four helices, M1–M4, overall forming a cluster of 20 TM helices. The pore lining helix, M2, is tilted radially inward toward the central axis up to the middle of the membrane. The outer helices (M1, M3, and M4) tilt both radially toward and tangentially around the central 5-fold axis.[264] Comparison of the ligand-free nAChR and ligand-bound AChBP structures suggests that ACh binding induces a local structural rearrangement (closure of two loops around ACh) to convert the α-subunits to their open (relaxed, R) state, which cooperatively triggers a transient opening of the channel pore at a distance of about 40 Å, thus allowing cations, particularly Na^+ and K^+, to pass through.

Several models have been proposed for elucidating the gating mechanism of nAChR.[267–272] NMAs performed by different groups for the complete structure of nAChR[270] and for the EC-TM domains of the homopentameric α_7 nAChR models based on the nAChR and AChBP structures[271,272] invariably showed that the lowest frequency mode is a concerted *quaternary twist* with counter-rotations of the EC and CP domains around the 5-fold symmetry axis (Figure

14A). Like all vibrational modes, this global mode gives rise to two sets of conformers, corresponding to positive and negative movements along the mode axis, manifested as opposite torsions in this case. Of these two sets, one is found to induce an opening in the TM channel of nAChR: the counterclockwise torsional rotation of the TM domain accompanied by clockwise rotation of the EC domain when viewed from the CP region. As can be seen in Figure 14B, the five M2 helices lining the pore are displaced slightly away from the center during this particular quaternary twisting. The calculation of the pore size profile along the TM channel (using HOLE273) shows that a relatively small (up to ~3Å) increase in diameter is induced in the constriction zone, the original value of which is 5.7 Å in the known structure. The diameter of the first hydration shell of a monovalent cation is typically around 8 Å. This small opening of the pore induced by the global mode is thus expected to enable the passage of hydrated cations.271

An increase in the pore radius by ~1.5 Å has indeed been suggested by MD and Brownian dynamics simulations to be sufficient to raise the computed conductance to ~22 pS—a value comparable to the experimental measurements for the open channel.[274] The above results from NMAs (including those obtained with ENMs) support the view that small but concerted rearrangements of the M2 helices lining the pore readily allow for an expansion of this size in the pore, thus providing an efficient gating mechanism. Concerted rigid-body motions of M2 helices were inferred by Unwin from early comparisons of the original structures at various resolutions.[275] Grosman and co-workers made extensive single-channel electrophysiological measurements to analyze the change in the microenvironment of the helices M1, M2, and M3 between the open and closed forms of the channel.276,277 Mainly, they examined the position-dependent proton transfers (or pK_a shifts) for ionizable residues that have been engineered in the inner faces of these helices. These experiments led them to conclude that nAChR pore dilation involved subtle rearrangements, if any, of these three helices.[276,277] Notably, the twisting mode predicted by the NMA does not necessarily implicate any significant change in the orientation of the M2 helix side chains with respect to the channel lumen but small rotations of about ~15° that presumably induce minimal changes in the exposure of side chains, which may explain the experimental observations. The changes induced by the NMA-predicted quaternary twisting mode, in the exposure of M2 residues' side chains to the central pore, were indeed pointed out by Changeux and co-workers to be compatible with the experimental data from Grosman and co-workers.[271]

The global *twisting-to-open* motion of nAChR resembles those observed in other multimeric ion channels, discussed above. The collective modes of the M2 bundle (pore-lining helices) predicted by NMA are also observed in PCA of MD simulation trajectories.[268] Conventional MD simulations of 30 ns for nAChR embedded in an explicit lipid bilayer also indicate269 the concerted rotations of M1 and M2 helices accompanying the shrinking of the ACh binding pocket, and the open–close transition of the structure can be driven by introducing a torsional rotation around the pore axis in steered MD. The accord between NMA results for the nAChR, in the absence of a lipid environment, and MD trajectories conducted in explicit water and lipid bilayer corroborates the dominant role of the membrane proteins' intrinsic features in defining the movements that facilitate essential functions such as gating.

In addition to gating, the mechanism of signal transduction from the ACh binding site to the pore, which presumably triggers the channel gating, has been a topic widely studied by both experiments and computations. In particular, the allosteric roles of individual residues and loops potentially involved in communicating agonist binding have been examined. Sine and co-workers identified, for example, the loops at the interface of the EC and TM domains that are required to couple the ligand-binding and pore domains in the serotonin type-3A receptor278 and identified the key residues within these loops, which signal agonist binding.[279] Single-molecule measurements of open-like vs closed-like propensities (in terms of Φ-

values) of individual residues at the transition state of the receptor[280,281] suggested a Brownian cascade of domain motions, whereby the transmitter binding domain assumes an open state and the M2 helices move toward the open state in discrete steps. Such a sequential cascade of discrete changes is not compatible with the all-or-none MWC-type allosteric motions predicted by NMA. The normal mode motions in the low frequency regime are smooth and concerted movements that simultaneously engage both the ACh binding and pore domains, rather than gradually progressing from one site to another. More recently, Sine and co-workers showed that the closed-to-open transition of the receptor involves two primed closed states independent of agonist binding.282 The primed closed states elicit short- or long-lived openings. Structural mapping of these states eludes computational studies due to the limitations in the resolution of the structures and those of the computational methods themselves.

The recently resolved X-ray structures of two bacterial homopentameric ligand-gated ion channels shed further light into pore opening/closing mechanisms. These are the closed state structure of the *Erwinia chrysanthemi* ligand-gated ion channel (ELIC)[283] and two open-state structures of the *Gloebacter* V*iolaceus* ligand-gated ion channel (GLIC).284,285 These structures do not include the CP helical bundle but bear EC and TM domains comparable in size and fold to their counterparts in nAChR. In particular, their EC domain superimposes closely with AChBP and with the EC domain of nAChR, except for a missing α-helix. The most striking difference between the ELIC and nAChR structures is at their pore domain: the EC half of the ELIC pore is occluded with Phe246 and Leu239 side chains that narrow down the pore diameter to less than 1 Å, while the remaining CP half is wide open (diameter of 6 Å). The two GLIC structures, on the other hand, are in the open state, being crystallized in the presence of an activating ligand proton. Figure 14C compares the ELIC and GLIC structures after their optimal superimposition. In addition to a symmetric tilt of the pore forming helices, the most visible difference is a quaternary twist similar to that observed in nAChR. Bocquet et al.[284] reported that the lowest ENM mode, a quaternary twist of the two domains, explains 29% of the structural difference between the cores of the structures. Overall, these structural data are consistent with a model of pore opening involving a quaternary twist and tertiary deformation.[284]

3.2.2. Rhodopsin—G-protein coupled receptors constitute one of the largest protein superfamilies in the human genome, with more than 800 members. Among the five families that form this superfamily, the rhodopsin family is the largest, with 701 members.[286] All GPCRs share a common architecture of seven TM α-helices (H1–H7) (Figure 7E). They transmit EC signals to the CP region via allosteric movements of TM helices. The resulting changes in the CP surface facilitate G-protein binding and activation, which, in turn, triggers a cascade of cellular responses.[287,288]

The vast majority of the structure-based computations for GPCRs have been done using the bovine rhodopsin structure, originally resolved by Palczewski and co-workers.[289] In addition to the bundle of seven TM helices, referred to as opsin, the structure contains an 11-*cis*-retinal (chromophore) deeply embedded in the core (Figure 15A). The EC domain consists of the N-terminus and three interhelical loops EC1–EC3; the CP domain contains three interhelical loops CL1–CL3 connecting respective pairs of helices H1–H2, H3–H4, and H5–H6, and a C-terminal helix H8 that runs parallel to the membrane. The EC domain contains a *β*-sheet, which serves as a lid to the chromophore binding pocket, stabilized by a highly conserved disulfide bond between Cys110 and Cys187. The retinal, covalently bound Lys296 on H7, undergoes a *cis*/*trans* isomerization upon light activation. This gives rise to a local conformational strain that propagates through the concerted rearrangement of the TM helical bundle to the CP domain, inducing an opening at the conserved D(E)RY motif, which is recognized by the G-protein (Figure 15B). The active form, metarhodopsin II, is reached after a series of

photointermediates. It binds the heterotrimeric G-protein, transducin, and interacts with several other signaling proteins.

Recent years have witnessed a remarkable progress in the number of newly solved GPCR structures.[290] Comparison of the structures of bovine opsin in its G-protein-interacting form (referred to as opsin*)[291] and rhodopsin shows, for example, an outward tilt of 6Å in TM6, and pairing of TM5 to TM6, in agreement with aforementioned experimental data and the above computational model proposed for metarhodopsin II. Comparison of the ligand-free opsin292 and opsin*, on the other hand, shows little structural difference, suggesting that the opsin conformational population is shifted toward the activated state in the absence of retinal and G-protein. In addition to opsin structures, the structures of four other unique GPCRs were recently determined,290 including the β_2 adrenergic receptor (β_2AR) structure by the Kobilka lab, an additional β_2AR structure that pointed to the stabilizing role of cholesterol,[290] the structures of an antagonist-bound A_{2A} adenosine receptor,[293] turkey β1-adrenergic receptor,[294] and squid rhodopsin.[295] All together, these structures provide important information on the divergent EC regions, differences in the ligand binding sites, and convergent features of the TM domains.

The type and extent of conformational changes undergone upon activation of rhodopsin have been extensively examined by various experiments[128,134,150–155,161,296–298] and computations.[85,180,216,299–306] GNM and ANM studies[216,300] show that the global mode is controlled by a broad hinge-bending region that includes the chromophore binding pocket and a number of highly constrained conserved residues in the close neighborhood (e.g., E113, T118, E122 on H3, F261 and W265 on H5, Y268 on H6, and C187 on β4) such that the structural changes locally induced upon the isomeric transition of the *cis*-retinal are efficiently propagated through cooperative rigid-body movements of the TM helices, toward both the CP and EC regions. An effect of these cooperative movements is opening the CP ends of the TM helices 3, 4, and 6, thus exposing the ERY motif at the G-protein binding site (Figure 15C). A model for the Meta II state has been proposed[300] by analyzing the lowest ANM modes in conjunction with experimental data. The model was shown to correctly predict 93% of the experimentally observed effects in 119 rhodopsin mutants for which the decay rates and misfolding data have been reported, including a systematic analysis of Cys → Ser replacements.[300]

With the elucidation of a large number of structures, we are now in a position to examine more closely the correlation between the experimentally observed structural differences and theoretically predicted conformational changes. We performed a PCA of currently available rhodopsin and opsin structures and compared the resulting PC modes to ANM modes. Our data set includes 16 structures, comprised of 14 rhodopsin and two opsin X-ray structures. Out of N = 348 residues, 312 are commonly resolved in the data set of examined structures, excluding the segments 230–240 on CL3, 311–313 between H7 and H8, and 327–348 at the C-terminus. The distribution of the structures along the first two principal modes is shown in Figure 16 panel A. These two modes contribute about 62% and 12%, respectively, to the structural variability in the data set. The PCA clearly separates the structures into two clusters along the first principal axis. These two clusters may be viewed, in a sense, as the two substates illustrated in Figure 2. Notably, the first cluster includes all the 14 rhodopsin structures in the inactive (sub)state and the second, two opsin conformations in the putative active (sub)state. Mode 1 therefore unambiguously distinguishes between these two substates, representative conformations of which are displayed in Figure 16 panel B. The second principal mode, on the other hand, further disperses the structures within the first cluster. This mode essentially refers to the changes in loop conformations and termini orientations. These can be viewed as the microstates in the inactive substate.

Rhodopsin thus provides an excellent example of how functional modes can be determined through PCA and NMA. One utility of PCA is to provide us with a simple organization of the ensemble of conformations accessible to a given protein, and this use will become increasingly valuable with the growth in PDB structures for the same protein, hence the development of PCA servers to perform such tasks.[44] The second utility is to assist in our assessment of the dominant changes in structure, which is usually described by the 3*N*-dimensional PCA mode 1, p_1, and the corresponding amplitude of motion scales with $\sigma_1^{1/2}$ (see eq 7). Panel C in Figure 16 illustrates how the rhodopsin conformation (red) is closely reproduced upon reconfiguring the opsin structure along p_1. Comparison of the range of the principal axes 1 and 2 in Figure 16 shows that the size of motions along p_1 is at least twice as large as that along p_2. Third, and most importantly, the principal modes may be directly compared with those predicted by NMA. The PCA modes are exclusively based on experimental data for an ensemble of structures, while ANM modes are predicted by the theory for a single structure. Comparison of the two sets can help benchmark the computational predictions, provided that the experimental data set represents a more or less complete ensemble (see, for example, the study performed by the Jernigan lab for HIV-1 protease190), or consolidate the results, given that both sets involve approximations. In the present case, the set of PDB structures is far from complete. Yet, ANM calculations performed for the two representative structures (labeled) from each cluster showed that p_1 exhibits a cumulative overlap of 0.79, with the first 20 ANM modes intrinsically accessible to opsin, and a cumulative overlap of 0.74, with the first 20 ANM modes accessible to rhodopsin. Thus, 2% of ANM modes in the low frequency regime provide a reasonable description of the change observed experimentally. The reconfiguration predicted by moving the opsin structure along these ANM modes is shown in panel D. These results again confirm the view that the relative movements of the TM helices 5 and 6 observed upon light activation are intrinsic properties encoded in the rhodopsin architecture.

3.3. Transporters

Transporters are generally active carriers. They require an energy-producing process to translocate a substrate against its concentration gradient. Secondary active transporters take advantage of the movement of a solute down a concentration gradient, so as to translocate another substrate across the membrane. Glutamate transporter, discussed below, is an example of such a transporter, where the uphill translocation of glutamate is coupled to downhill Na^+ transport.

3.3.1. Glutamate Transporters—The concentration of glutamate in the EC space may increase by 10^3–10^4 fold during periods of synaptic activation, and it is critical to have a mechanism in place to clear the excess glutamate, which, otherwise, may have neurotoxic effects. Glutamate transporters clear the excess glutamate molecules. They are located on neurons and glia (Figure 17). Their precise functioning (glutamate uptake and reuptake) is essential not only to protect against excitotoxicity316,317 but also to regulate glutamatergic signal transmission by preventing sustained activation and desensitization of ionotropic receptors and modulating the activation of metabotropic receptors.318 Members of this family, including the human excitatory amino acid transporters (hEAATs), utilize Na^+-derived electrochemical gradients to transport glutamate, hence their classification as the family of dicarboxylate/amino acid: cation symporters,[319] also referred to as the glutamate transporter family.[320] Notably, these transporters also function as chloride channels.[321,322]

The first member structurally resolved in this family is an archaeal aspartate transporter, Glt_{Ph}, from *Pyrococcus horikoshii*.[323,324] Glt_{Ph} provided for the first time a structural model for gaining insights into the molecular basis of glutamate transport by the human orthologs, hEAAT1–5.[325] The top view (from the EC side) and side view of Glt_{Ph} can be viewed in the respective Figures 18A and 19B. As can be seen, Glt_{Ph} is a homotrimer, the three monomers

of which are arranged cylindrically to form a bowl-shaped basin of ~50 Å diameter and ~30 Å depth toward the EC surface at the center of the protein. Each subunit is composed of two domains: the N-terminal domain consists of TM1–TM6 (gray in Figure 18A), and the C-terminal core is comprised of the helices TM7 and TM8, and helical hairpins HP1 and HP2 (colored in Figure 18A).[324]

We note that TM7 has an unusual structure, with two helical segments connected by a partially unwound conserved motif $N_{310}MDGT_{314}$.324 This motif is in a sense "frustrated", containing potentially hydrogen-bond-forming groups that lack partners, hence the role of this motif in binding the substrate (confirmed in simulations326) and the clustering of several conserved residues near this region (Figure 18A). In other words, the irregularity of this helix at its central portion is actually functional. We note that this region is located exactly near the tips of the two loops HP1 and HP2. Another region important in substrate recognition is indeed the serine-rich tip of the HP1 loop, which is in van der Waals contact with the tip of HP2. The amphipathic TM8 helix also shows irregularities, which have been observed in simulations[326] to be involved in substrate binding and channeling.[328] HP8 has indeed been proposed to form a portion of the transport pathway.[320] Intersubunit contacts are confined to the N-terminal domain, suggesting that the three core domains function independently of each other, as also suggested by experiments.[327]

We recently examined the substrate recognition and binding events of Glt_{Ph} by MD runs of tens of nanoseconds.[326] Our simulations clearly showed that the HP2 hairpin acts as an "EC gate", in accord with the mechanisms inferred[324] from structural data. The fluctuations of this gate between its open and closed substates occur within the time scale of nanoseconds. Figure 18B illustrates the sequence of events observed[326] in a typical run, starting from diffusion of the substrate toward the partially open "gate" and continuing with the recognition of conserved glycines (G354 and G357) at the HP2 tip, gradual insertion into the region between HP1 and HP2, near a conserved 3-Ser motif on HP1, and subsequent stabilization at a site near the unwound portion of TM7. The latter site and geometry are in close agreement with the binding pose experimentally observed[323] for aspartate in the high resolution structure of Glt_{Ph}.

These conventional MD simulations thus provided ample information on the early recognition and binding events. However, no substrate translocation could be observed in these simulations. Instead, we performed nonequilibrium steered MD (SMD) simulations,[328] which helped elucidate two substrate translocation pathways, one of which is more readily accessible, along with the key interactions and energy barriers encountered during the translocation. The SMD thus provided information on "slower" events, which would otherwise be inaccessible via classical MD.

Notably, both substrate binding and substrate translocation events examined in these two sets of simulations are local events that involve the core regions in each subunit, and they appeared to occur independently in the three subunits, consistent with experimental observations. There is, however, yet another dimension: the movements that cooperatively engage all three subunits, which may perhaps explain the *raison d'être* for the functioning of glutamate transporter as a trimer, rather than as three monomeric proteins. ANM analysis of the global dynamics of Glt_{Ph} yields the movements depicted in Figure 18C as the first nondegenerate mode (mode 3), along with a doubly degenerate asymmetric stretching/contraction mode accessible to the transporter. In mode 3, the three subunits undergo concerted opening/closing movements to alternately expose and cover the central basin. The diagrams in the figure are color coded, with the red portions corresponding to the most mobile regions. The largest movements are undergone in this case by the N-terminal domain residues that are exposed to the EC region and the aqueous basin. It is interesting to note the possible occurrence of intersubunit contacts between the EC-exposed portions of TM8 and HP2 during the concerted

movements of the three subunits in this mode. The diagram and arrows in Figure 19 panel B illustrate the same movement, viewed from the side; and Figure 19A displays the corresponding mobility profile (normalized distribution of residue square displacements driven by this mode). G144 and T182 exhibit minimal mobilities. These two residues are located at the base of the aqueous basin on the EC and IC sides, respectively, and are proposed to play a role in modulating the concerted motion of the subunits. The base of the EC vestibule indeed remains rigid and immobile during these movements, which may be a requirement to maintain the integrity of the trimer. As to the peaks in the mobility profile, we note a number of histidines exposed to the EC region. It is interesting to note, in this context, that Vandenberg and co-workers observed that Zn^{2+} ions inhibited the anion conductance of EAAT4 and attributed this inhibition to the binding of Zn^{2+} to *His146* and *His154* conserved in EAAT1, -2, -4, and -5.[329] Interestingly, these two histidines lie very close, both sequentially and spatially, to the top-ranking residue (His114 in Glt_{Ph} or *Lys152* in EAAT1) in the global mode profile, suggesting that the peak residues observed here could possibly serve as "sensors" for capturing negatively charged substrates.

3.3.2. ATP Binding Cassette Transporter BtuCD—ATP binding cassette transporters mediate the transport of various substrates, including ions, drugs, lipid molecules, and small proteins, across the membranes via an ATP-dependent mechanism.[330,331] BtuCD is a member of the family of ABC transporters that transports vitamin B_{12} in *E. coli*, and it has been investigated by both structural and computational studies.

The BtuCD complex, like most ABC transporters, consists of four subunits, arranged as two homodimers: the TM dimer, BtuC, that forms the specific substrate translocation pathway, consisting of 20 TM helices; and the CP dimer, BtuD, composed of two nucleotide-binding domains (NBDs) where the ATP binding and hydrolysis activities take place.[332–334] The TM dimers of ABC transporters usually exhibit little sequence similarity, with their sequence being specific to the particular substrate that they recognize and translocate. The NBDs, on the other hand, are highly conserved, sequentially and structurally. They share the common motifs Walker A and Walker B typical of many ATP-binding domains, as well as a signature sequence, also called the C motif, that completes the ATP binding site at the interface between the two NBDs.

Examination of the crystal structure of BtuCD[332] and MD simulations performed for BtuCD[335] and MalK (an ABC transporter with similar architecture[336]) suggests that ATP binding to the BtuD dimer triggers conformational changes that propagate to BtuC. ATP binding induces a closing at the interface between the two NBDs of the BtuD dimer and stabilizes the interactions (or reduces the fluctuations) at the interface, as confirmed by MD simulations[337] of the dimer. However, the extent to which these structural changes occur in the presence of the BtuC dimer has been questioned, given that the MD simulations for the tetramer did not exhibit the decrease in the inter-NBD distance observed upon ATP binding in the BtuD dimer.[337] A related question of interest is the role of BtuF, the protein that delivers the vitamin B_{12} to the periplasmic entrance of the BtuCD, on the dynamics of the transporter.

NMA of BtuCD dynamics was recently performed by Ma and co-workers toward elucidating the intrinsic dynamics of the transporter.[338] In this study, the lowest frequency modes accessible to the two dimers, BtuC and BtuD, were analyzed separately with ANM, as were those accessible to the tetrameric BtuCD structure (of ~1100 residues). This analysis demonstrated that the lowest mode accessible to the tetramer, ANM mode 1, is a highly cooperative motion that involves the reverse rotations of the two halves of the BtuCD transporter, as illustrated in Figure 20, top panel. Interestingly, the relative movements of the two TM domains in this mode correlate (with an overlap of 0.83, after removing the rigid-body contributions) with the mode 1 accessible to the BtuC dimer alone,[338] suggesting that the

intrinsic dynamics of BtuC strongly affects this global mode of the tetramer. This mode was also shown to be insensitive to BtuF capping.338

The mechanism of BtuCD mode 1 is more clearly visualized from the side (middle diagram in Figure 20 top panel, obtained by rotating the diagram on the left by 90° around the vertical 2-fold symmetry axis). This view shows that monomers 2 and 4 (belonging to the respective dimers BtuC and BtuD; see labels in panel A) rotate almost rigidly together, while monomers 1 and 3 undergo an opposite rotation with the pivot located near the CP gate of the translocation pathway.[338] As a result, this movement induces a simultaneous shear opening between the pairs of NBDs (3 and 4) or their ATP binding sites and between the pairs of TM domains (1 and 2) near the EC ends of the TM helices, while the CP gate undergoes minimal, if any, displacement (see also the color-coded diagram on the left). Thus, closing of NBDs upon ATP binding is accompanied by a simultaneous closing of the EC pore, which prior to this allosteric effect was sufficiently large to accommodate vitamin B_{12} entry to the translocation pathway. Thus, Weng et al. proposed that the vitamin B_{12} molecule is trapped into the periplasmic cavity rather than transported to the CP region, upon ATP binding.[338] Weng et al. further proposed that the conformational change required for substrate translocation and the opening of the CP end of the pore is associated with ATP hydrolysis (rather than ATP binding). In particular, mode 3 has been proposed to contribute largely to the conformational change powered upon ATP hydrolysis. As can be seen in the bottom panel of Figure 20, this mode induces an opening at the pore region of the BtuC dimer. BtuF capping appears to restrict these movements, while another mode (mode 7, not shown) has been pointed out to effectively enable substrate transport. Modes 3 and 7 of the tetramer have been shown to be similar in shape to the lowest two modes favored by the BtuD dimer structure,[338] which supports the significance of these intrinsically accessible modes in mediating the ATP-dependent coupling of the dimers.

4. Conclusion

Over the years, many techniques have been developed to tackle an ever broadening range of problems using NMA in general and ENMs in particular. The linearity of the theory endows it with considerable flexibility, and the clever applications of matrix algebra to NMA have expanded its utility. In particular, ANM and its extensions have been of great use in studying dynamic phenomena that exceed the time or length scales of MD, such as investigations of Megadalton-scale structures' dynamics,[64,81,207,232,233,339] exploring the CG transition pathways,[340–346] and studying the effects of crystal packing on protein dynamics.[194,347–350] Other studies have taken advantage of the computational efficiency of the GNM/ANM to perform serial analyses of large data sets and gain insights into design principles. An example is the colocalization of global hinge sites and catalytic sites in enzymes, which appears to be a mechanism of efficiently coordinating the mechanical and chemical activities of the protein. 32 Another example is the intrinsic ability of the proteins in the unbound form to undergo structural changes that are stabilized upon substrate binding.21,30,[351] We presented several applications to membrane proteins in section 3. Below, we present an overview of insights into mechanisms and principles of functional dynamics gained from ENM-based studies (section 4.1) and recent extensions that are anticipated to be exploited and further developed in future studies (section 4.2).

4.1. Robustness and Functionality of Global Modes

4.1.1. Robustness of Global Modes: A Requirement in Evolutionary Selection of Structures—Designable protein *structures* are usually referred to as structures that are the lowest energy conformer for a multitude of sequences; that is, these structures can usually tolerate sequence substitutions with minimal change in their overall fold. This type of

insensitivity to sequence variations is what makes a stable structure, in a sense. But stability does not necessarily imply functional aptitude.

Function, on the contrary, requires a well-defined flexibility and conformational malleability, within a coarse-grained view of the global energy minimum, perhaps evidenced by substates that are accessible via small energy barriers. In the same way that stable structures are those which are insensitive to sequence variations, one can think of functional proteins as those whose dynamics are insensitive to structural details. Indeed, the success of ENM-based NMAs presumably originates from the insensitivity of global modes to structural and energetic details.

The observed robustness of global modes may reflect an evolutionary pressure. Stable structures are those mapped onto by *many* sequences, according to the designability principle set forth by Wingreen and co-workers.[352] Functional structures, on the other hand, are proposed to be those that intrinsically favor the global modes that facilitate/accommodate biological functions such as substrate binding, translocation, or gating by membrane proteins; the global modes are in this case favored, or mapped onto, by the overall architecture despite minor changes/perturbations in structure.

4.1.2. Toward Gaining Insights into Functional Dynamics of Membrane Proteins

—Biomolecular dynamics is a complex process. In particular, the transitions between conformational states separated by high energy barriers, such as the folding of proteins, continue to pose a challenging problem—except for small proteins where some success has been recently achieved. The transitions between microstates within a global energy well, or between substates separated by relatively low energy barriers, on the other hand, appear to be a more tractable problem, with the development and applicability of elastic network models and PCA-based methods. The rapidly increasing structural data now permit us to test and improve these coarse-grained models and methods. In the present review, we illustrated the recent applications to membrane proteins, as a group of proteins that are extremely important from biological and pharmaceutical points of view.

These studies provide us with insights into the collective mechanisms of motions preferentially accessed by membrane proteins. A striking observation is the occurrence of a global twisting as a mechanism of pore opening or ligand gating in many membrane proteins. The “twist-to-open” mechanism instrumental in the gating function of most of the membrane proteins discussed here suggests a common mechanism of pore-opening when the pore architecture exhibits a cylindrical symmetry with funnel-like organization of a bundle of helices. Another observation is the high cooperativity of the motions, which becomes even more pronounced by the structural symmetry or multimerization. In this respect, nAchR presents a unique structure, being a heteropentamer. Yet, the dominant mechanism conducive to channel opening is again observed to be a global twist. It is also interesting to note that the transition between the closed and open forms has been observed in many applications, to be realized by a small subset of modes at the low frequency regime, and among them, nondegenerate modes usually provide the most productive paths leading to functional substates in the case of structurally symmetric multimers. The passage between the open and closed forms of the MscL achieved by a few nondegenerate modes is a typical example (Figure 13).

ENM-based NMAs not only provide insights into the most easily accessible movements of quaternary structures but also point to sites that may play a critical role in mediating or propagating allosteric signals. In the same way as there are particular amino acids whose substitution may be deleterious (conserved residues) to stability, there are particular sites on the structure whose perturbation could impair the global dynamics (e.g., hinge sites in the global modes). We have learned that these sites are utilized by proteins to elicit cooperative responses, e.g., ligand binding pockets that efficiently transmit allosteric signals, especially if fueled by

the energy released by an exothermic reaction (e.g., ATP hydrolysis) in the vicinity. These sites are referred to as mechanically critical sites[32] or sites with a high allosteric potential.[238,239] Not surprisingly, more and more structures show us that active sites, the drug binding sites, or residues that are known to mediate allosteric effects, or ATP binding sites, coincide with, or closely neighbor, such mechanically critical sites. It is clear that an improved understanding of the structural basis for allosteric and chemical communications in these proteins will assist in the rational discovery of drugs against the various channelopathies or signaling diseases.

4.1.3. Many Functional Motions of Membrane Proteins Are Intrinsic to Their 3D Structure, Independent of Membrane Environment—Evidence for the dominance of intrinsic dynamics in defining certain collective motions and/or allosteric mechanisms of membrane proteins (such as gating or signaling), independent of the membrane environment, is provided by the applications presented in section 3.

We began with gramicidin A, for example, in section 3.1.1. The calculations performed by Miloshevsky and Jordan unambiguously demonstrated the equivalence of the NMA results from ENMs and those from full atomic models in the presence of explicit lipid and water molecules subject to the CHARMM22 force field.192 In particular, the gating mechanism (counter-rotation of the two helices) was concluded to be an inherent property of the GA architecture, independent of surrounding lipid and water molecules. Likewise, the ANM calculations performed for a series of potassium channels by Shrivastava and Bahar218 yielded results (cooperative rotational/twisting motions of M2 helices to induce pore opening) in excellent agreement with SDSL EPR data from Perozo's lab[157,158] and confirmed by recent structural data (kink region) determined for a newly resolved cation channel (NaK),251 again showing that the rigorous consideration of the native contact topology permits us to predict global movements relevant to function regardless of the potential perturbation of the membrane environment. In the case of nAChR, the quaternary twist model from ANM studies[270–272] not only agrees with the mechanisms inferred from MD and Brownian dynamics simulations[268,269,274] but also compares favorably with the newly elucidated closed and open structures of the ligand-gated ion channels GLIC[284,285] and ELIC.[283] The global movements of the archaeal aspartate transporter Glt_{Ph}, on the other hand, drive the cooperative opening/closing of the three subunits around the central aqueous basin, which were not observed in tens-of-nanoseconds simulations. The computations with Glt_{Ph} using MD,326 steered MD, 328 and ANM indeed provide a nice example of the complementarity of results and the utility of exploiting multiple scale computations. An even better approach is to develop integrated MD-ANM approaches, like the ANM-steered MD applied to rhodopsin,[85] which simultaneously provided access to global movements while viewing atomic interactions and rearrangements. A striking observation in this case is the close correspondence between ANM modes and the dominant modes derived from the PCA of 16 X-ray structures resolved for rhodopsin in different forms. The fact that a small subset (2%) of ANM modes in the low frequency regime yields an overlap of ×0.75 with the principal modes of deformations derived from experimental data again lends strong support to the physical and biological significance of ANM modes, which are based exclusively on the protein structure, independent of membrane environment.

The mechanisms of collective movements essential to certain functions such as gating or allosteric signaling thus appear to be intrinsic to protein structure, in accord with ENM-based predictions exclusively based on the inter-residue contact topology of the membrane protein. However, the function of membrane proteins involves many other specific and subtle interactions that cannot be studied by CG models and NMA, such as the selection of particular ions at the selectivity filter, recognition and binding of substrate by specific interactions, and the assistance of substrate translocation by cotransported ions. Moreover, ENM-based NMA

cannot provide information on the absolute time scales of the movements either, due to lack of a proper consideration of the frictional drag or other environmental factors that may affect the relative frequencies or probabilistic occurrence of different modes of motions. Essentially, the ENM-based studies provide information on the "accessible" most cooperative movements that are selected/recruited functions. However, there also exist several accessible, energetically favorable movements at local scales, including side chain isomerizations or specific reorientations of polar groups, which require full atomistic and, in some cases, even quantum mechanical calculations. Adding to the complexity is the potential coupling between local events and global movements, hence the need for developing multiscale methodologies that take advantage of the capabilities of both MD and NMA.

4.1.4. Entropic vs Enthalpic Effects, or Geometry vs Specificity—It is important to note that ENM-based approaches are based on purely geometric considerations such as inter-residue contact topology or overall shape/architecture of the examined structure. As a result, the predicted movements are those which are entropically favored, as originally proposed for polymeric networks,[69,70] and do not contain contributions from specific (enthalpic) interactions. ENM-based approaches are therefore useful to the extent that geometry or topology plays a dominant role in the process being explored. The machinery of supramolecular systems could be a prime example for processes dominated by collective mechanics, rather than specific/local chemical events. However, in many applications it may be important, if not indispensable, to invoke both effects and adopt hybrid or multiscale approaches. Recent years have indeed seen a large number of studies in that direction, which have been partly reviewed here. Overall, it is important to interpret the ENM-predicted dynamics as one aspect of the complex mechanochemical process, mainly that intrinsically preferred by the particular architecture, in the absence of perturbations from specific intra- and intermolecular effects.

4.2. Extensions of Coarse-Grained NMA and Future Directions

4.2.1. Hybrid Methods That Integrate CG NMA and MD—As discussed above, MD simulations realistically explore, in the presence of explicit solvent and/or membrane, events on the nanosecond time scale for biomolecules of a few hundreds of residues, while their application to larger systems (e.g., multimeric proteins) and longer processes (e.g., of the order of microseconds or slower) suffers from sampling inaccuracies. NMA with ENMs provide an understanding of the global movements of Megadaltons systems, but at the cost of loosing accuracy and specificity at the local scale. The above-described applications of MD and NMA-methods to membrane proteins nicely illustrate the capabilities and limitations of the two sets of computations. For example, the CG modeling and the atomistic simulations of inward rectifying potassium channel Kir3.4.1[353] provided similar pictures of the overall dynamics of the ligand-binding domain, suggesting dimer-of-dimers motion as an intrinsic property of the CP domain of this K^+ channel. Thus, combining the data from these alternative computational approaches may help consolidate the inferred mechanisms, if a consistent behavior is captured. Likewise, the PCA of MD trajectories generated for BtuCD in a lipid bilayer supports the hypothesis that ATP-binding drives closure of the nucleotide binding domains in BtuD, while the apo state of BtuD randomly switches between open and closed substates,[337] consistent with ANM analysis of the same dimer. It is not generally sufficient, however, to perform and compare two independent sets of computations, such as MD and CG NMA, to make inferences on multiscale dynamics. Instead, there is a need to develop hybrid methodologies or more integrated approaches that exploit the complementary utilities of the two methods and take account of possible couplings between these different scale events.

A new protocol that steers MD along ANM modes has been recently developed to this aim and used to study of the conformational changes associated with photoactivation signal transmission in rhodopsin.[85] In this method, global conformational changes that are not

accessible via conventional MD trajectories can be sampled, while motions and interactions at atomic scale can be observed in the presence of explicit solvent and lipid bilayer. Two stable regions were identified by this method for rhodopsin, one clustered at the chromophore and the second at the CP end of the TMs 1, 2, and 7 (Figure 21). These simulations elucidate the redistribution of the interactions between the retinal and its neighboring residues on H3–H6, induced upon *cis* → *trans* isomerization of retinal. Eleven of the 16 residues identified to participate in the central hinge region near the retinal have been tested by experiments and confirmed to play a critical role in stabilizing the activated state. Furthermore, these simulations draw attention to the possible role of water molecules in coordinating the interactions between conserved residues at the CP ends of the helices H1, H2, and H7, illustrated in Figure 21.

Another study in the same spirit is the examination of the allosteric changes in the conformation of BtuCD TMDs. These movements that appear to be induced upon ATP binding to the BtuD dimer have been examined by Tieleman and co-workers using the perturbed ENM (see subsection 2.3.4) and biased MD simulations.[240] The results support the MalK model for the transport mechanism; that is, a closure of the nucleotide binding domain upon ATP binding is predicted, which results in closing of the TMD toward the CP side while inducing an opening toward the periplasmic side.

4.2.2. Docking and NMA in Drug Discovery—Understanding the mechanism of interactions between the target protein and a small molecule inhibitor is of crucial importance in drug discovery.[102] Molecular docking is the primary computational tool to model these interactions[362] and screen compound libraries of small molecules with potential inhibitory/agonistic/antagonistic activities.[363] There are numerous successful applications of docking to membrane proteins. Predix Pharmaceuticals, for example, targeted five different GPCRs in *in silico* screens of commercially available libraries and identified 11 compounds per target, with an average hit rate of 17%.[364] In another study, Wang and co-workers targeted dopamine (D_3) receptors and identified four compounds that bind at 100 nM levels, with 60% hit rates.[365]

The ligand-selective conformational heterogeneity of GPCRs has been recognized, however, as a limiting factor in *in silico* efforts.[287,366] The binding site geometries of GPCRs differ, depending on the functionality and the potency of bound ligands.[367] Kinetic measurements and single molecule spectroscopy both reveal that the 7TM helix bundle samples distinguishable conformational states in the absence or presence of ligand, and the populations of these conformational states shift upon ligand binding.[368,369] State-of-art docking programs usually allow for only partial binding site flexibility limited to optimizing a small number of side-chain rotations or short loop conformers. Overlooking such conformational flexibilities hampers the success of *in silico* drug discovery.

Abagyan and co-workers made prominent contributions to developing algorithms and tools that take account of target protein conformational flexibility,[370,371] which have been successfully applied to GPCRs.[372] In particular, a ligand-steered homology modeling approach was developed, which uses existing ligands to shape and optimize the GPCRs binding site.[373] The idea therein is to start with hundreds of crude homology models as probable conformations of the target protein and then filter them based on their interaction energy with known ligands probed by flexible docking and on their ability to detect known ligands in virtual screening tests. The utility of this approach was demonstrated by its application to melanin-concentrating hormone receptor 1, where a 10-fold improvement over random high-throughput-screening was achieved and six novel antagonists were identified. In a similar recently published study,[374] β_2AR interactions with agonist/antagonist were examined upon generating multiple conformations of β_2AR. The models were reduced and further refined by flexible docking of selected agonists in the light of mutagenesis data to obtain models that

outperformed rhodopsin-based models. In accord with these findings, Kobilka and co-workers reported that rhodopsin-based homology models of β_2AR developed prior to β_2AR structure resolution were more similar to rhodopsin rather to β_2AR,[375] stipulating the need to consider more distinctive target conformations.

The generation of multiple conformations for the target protein emerges from the above and other studies[376–381] as an important component of computational tasks for modeling and simulating protein–inhibitor interactions. NMA with ENMs appears to be particularly suitable for generating backbone rearrangements. It suffices to have but one structure to generate a distribution of energetically favorable conformations in its neighborhood. Likewise, the method can be used to refine/broaden an existing collection of conformations.

Figure 22 illustrates three cases where such NMA-based generation of alternative conformers improved the performance of docking simulations. Panel A shows the results for cyclic AMP-dependent protein kinase.[382] A ligand binding loop in this protein is known to assume different conformations in the presence of different ligands. Alternative loop conformations favored by the structure were determined in this case by selecting from the ensemble of low-to-medium frequency modes those that specifically induce reorientations in this particular loop. The use of this ensemble in docking simulations was shown to improve the discrimination rate between binders and nonbinders.[382] Panel B refers to the study of matrix metalloproteinases inhibitors by Perahia and co-workers.[383] The global mode that directly affects the opening/closing of the ligand binding cavity was identified in this case to be the second lowest mode, and a set of conformations was generated by gradually reconfiguring the protein along that mode. Docking of inhibitors to the resulting target ensemble was shown to improve docking in all cases compared to docking to a single energy-minimized structure. Finally, May and Zacharias used NMA to improve protein–protein and ligand–DNA docking in a number of studies.[379,384,385] Panel C illustrates the application to cyclin-dependent kinase 2.384 An ensemble of ligands were randomly placed in the binding site of the protein in this case, and the protein–ligand interactions were optimized by deforming the protein along the lowest ten modes accessible to the structure.[384] The approach improved all poorly docked cases, at only a modest increase in computational cost. Two more studies from Perahia's lab support the utility of NMA-based modeling, one involving a protein that interacts with a membrane protein (CD47 receptor), and the other showing how collective motions (of a synthase) relate to its catalytic activity.[386,387]

4.2.3. Normal Modes for Structural Refinement—An obstacle to understanding the chemistry of biological molecules has always been the determination of their structures to atomic precision. Unless a protein readily crystallizes or is sufficiently small to produce a clean NMR signal, its structure cannot be determined to high resolution. In many cases, the structure can be predicted by homology modeling and then refined to locate the optimal conformation for the particular sequence. Feig and co-workers showed[388] that a good set of decoy structures against which to refine can be generated by distorting a homologous template along its slowest normal modes. Indeed, because these slowest modes indicate the easiest directions of motion, refinement using normal modes produces higher resolution structures than does refinement using other CG methods or MD.[388]

Normal modes are also used in refining electron microscopy (EM) structures. The idea is to use structural data available from X-ray crystallography or homology modeling for substructures (e.g., individual domains, subunits, etc.) and exploit their NMA-predicted alternative conformers to optimally fit cryo-EM data for the intact structure. Several methods have been developed for structural refinement using ENMs,[389–395] some of which are available as software packages.201,396 The basic technique in these approaches is to start with a known high-resolution structure and iteratively alter it along its normal modes, preferably the slow modes, until its structure agrees with the EM density map.

Vector quantization-based techniques have also been used to predict the collective motions of macromolecules from low-resolution structures.[397,398] The underlying idea therein is that an ENM constructed around the EM density map produces the same dynamics as an ENM constructed using the detailed structure that is represented by the density map. The map is divided into a set of discrete points that act as nodes in the ENM, and the global dynamics are calculated with NMA. This technique has been shown to predict motions in accord with experimentally observed fluctuations.[399]

Recently, the combination of MD and NMA results for DHFR complexed with nicotinamide adenine dinucleotide phosphate by the Perahia lab demonstrated that the inelastic neutron scattering spectrum may reflect proteins trapped in different conformations (at 120 K), in addition to the vibrational modes of different conformations, leading to inhomogeneous broadening of the spectrum.[400]

4.2.4. Exploring Allosteric Transitions in Large Biomolecular Systems—Proteins usually sample multiple substates, prompted by an external event, such as ligand binding or assembly with another protein, especially if an allosteric change in conformation is triggered. The details of the transition from one state to the other are in most cases only marginally understood; the transition likely does not follow a single linear trajectory but instead winds through a complex energy landscape. Nonetheless, these global transitions appear to proceed, or are at least initiated, via the collective global mode directions that are studied with NMA, and it has been shown in many applications that biomolecular structural transitions between functional substates are largely accounted for by a few slow modes. This observation puts ENMs among the primary tools for theoretical studies of transition pathways.

Since NMA is valid only in the local region surrounding a potential energy minimum, its application to nonequilibrium events such as conformational changes must be handled delicately. When studying the simplest case of a system with two stable conformation—call them “A” and “B”—it is assumed that each conformation resides at the bottom of a harmonic potential energy well and that the transition state is sufficiently close to both end points as to be within the range of the harmonic approximation about each conformation. The system can then transition smoothly from the harmonic surface surrounding A to the harmonic surface surrounding B.

An early technique for studying transition pathways using ENMs[340] involves constructing two ENMs for the initial conformation (Figure 23). The topology of the first ENM is determined by the contacts in state A, and the topology for the second is determined by the native contacts in state B. The system is initially modeled with the EN for state A only, and the transition proceeds by gradually reducing the effects of the state A EN while increasing the contribution of the state B EN. Throughout the transition, intermediate conformations are calculated by minimizing a cost function based on the instantaneous EN. This method has since been adapted to use rigid clusters or combinations of rigid clusters and pointlike beads.[341] A related method[342,343] involves constructing an ENM for the initial state only and then slowly perturbing the structure to satisfy known distance constraints from the final structure. A similar “nonlinear” elastic model, in which the modes are continuously modified in a series of small steps through the transition, was used to study the open to closed transition of adenylate kinase.[345] All three methods are capable of generating putative transition pathways with low computational cost.

Another technique, the plastic network model,[344] combines the potentials about states A and B into an analytical double-well potential. The transition states are defined as conformations on the cusp between states A and B. This model was also used to investigate the allosteric transition in adenylate kinase. A similar “mixed ENM” was used to investigate the helix-to-

sheet transition of the Arc repressor,[401] as well as transitions of kinesin and myosin.[237] The plastic network model was modified[402] to make the potential a double-well for all interactions. The resulting potential energy landscape is unlike the original but has multiple local extrema. This model, too, was used to investigate the open-to-closed transition of adenylate kinase, revealing an alternative transition pathway. Yet another method, the adaptive ANM of Yang et al.[346] gradually moves the structures from both end points along their respective slow modes until a common structure is reached at the assumed transition state. It is also worth noting that not all NMA-based methods of generating transition pathways are strictly analytical. The MC normal mode following method of Miloshevsky and Jordan[192] utilizes normal modes to guide simulations. An advantage of this technique is that it does not require two conformations but generates an approximate potential energy landscape from the initial conformation alone.

Besides the proteins discussed in this review article, there are also other membrane proteins where the normal modes predicted by CG NMA have been observed to correlate closely with experimental observations on gating mechanisms, for example acid-sensing ion channel (Asc1).[403] In this very recent study, modes 1 and 3 induce a twisting of the entire TM domain coupled to the motions of a β-turn, which in turn induces an opening of the channel pore. The predicted mechanism is consistent with mutagenesis and electrophysiological experiments. To quote the authors, “This result indicates that the structure of the closed, desensitized state of Asc 1 intrinsically tends to undergo a twisting motion to open the gate”. This study reinforces the point made in this review, namely, the “twisting-to-open” motion is a common mechanism for gating membrane proteins with an inherent cylindrical symmetry.

Acknowledgments

The authors benefited from Stephen White's Web site on membrane protein structure statistics, Dr. Yang Zheng's assistance in generating some of the ANM figures, and Dr. Basak Isin's useful comments on rhodopsin dynamics. Support by NIH Grants 5R01 GM086238-02 and 5R01 LM007994-06 and by PA Department of Health (No. 0317401) is gratefully acknowledged by I.B.

Biography

Ivet Bahar is the John K. Vries Founding Chair of the Department of Computational Biology, School of Medicine, University of Pittsburgh, and Director of the Joint Ph.D. Program in Computational Biology between Carnegie Mellon University and University of Pittsburgh. She served as an Assistant (1986–87), Associate (1987–93), and Full Professor (1993–2001) at the Chemical Engineering Department of Bogazici University, Istanbul, Turkey, before joining the University of Pittsburgh. Her research areas are biomolecular systems structure, dynamics, and interactions. She authored in over 170 papers in the areas of polymer chemistry, molecular biophysics, and computational structural biology. She has been an elected member of EMBO since 2000, a principal member of the Turkish Academy of Sciences, an elected member of the Biophysical Society Council, and an Executive Committee Member of the International Society of Quantum Biology and Pharmacology.

Timothy Lezon received his B.S. in physics from the University of Illinois at Urbana–Champaign in 1997 and his Ph.D. in physics from the Pennsylvania State University in 2007. His research interests include the application of statistical physics techniques to biological systems, particularly in the development of coarse-grained models for exploring protein structure and dynamics.

Ahmet Bakan received his undergraduate education at Koç University, Istanbul. He is currently pursuing his Ph.D. degree in computational biology at the University of Pittsburgh. His interests focus on protein–ligand interactions, protein dynamics, and the role of dynamics in recognition and binding events.

Dr. Indira Shrivastava is currently a Research Associate at the Department of Computational Biology at the University of Pittsburgh. She received her Ph.D. in Quantum Chemistry from the Department of Chemistry at University of Pune, India, in 1993. She did postdoctoral work at the Indian Institute of Science, Bangalore, EMBL, Heidelberg, and the Structural Biology Unit, at the University of Oxford, and was a visiting fellow at the National Institutes of Health, Maryland. Her research interests include molecular modeling and simulations of membrane proteins to explore functional dynamics of these proteins by an integrated approach of macroscopic, microscopic, and submicroscopic simulations. She dedicates this review article to her father, Mr. Harigovind Shrivastava, upon his 75th birthday.

6. References

1. Birshtein, TM.; Ptitsyn, OB. Conformations of Macromolecules. Interscience; New York: 1966.
2. Flory, PJ. Statistical Mechanics of Chain Molecules. Wiley, Interscience (2nd Edition published in 1989 by Butterworth-Heinemann Ltd); New York: 1969. p. 1-432.
3. Volkenstein, MV. Configurational Statistics of Polymer Chains. Interscience; New York: 1963.
4. Taketomi H, Ueda Y, Go N. Int. J. Pept. Protein Res 1975;7:445. [PubMed: 1201909]
5. Noguti T, Go N, Ooi T, Nishikawa K. Biochim. Biophys. Acta 1981;671:93. [PubMed: 6171304]
6. Go N. Annu. Rev. Biophys. Bioeng 1983;12:183. [PubMed: 6347038]
7. Levitt M, Warshel A. Nature 1975;253:694. [PubMed: 1167625]
8. Cooper A. Proc. Natl. Acad. Sci. U.S.A 1976;73:2740. [PubMed: 1066687]
9. Frauenfelder H, Sligar SG, Wolynes PG. Science 1991;254:1598. [PubMed: 1749933]
10. Frauenfelder H, Parak F, Young RD. Annu. Rev. Biophys. Biophys. Chem 1988;17:451. [PubMed: 3293595]
11. Frauenfelder H, Petsko GA, Tsernoglou D. Nature 1979;280:558. [PubMed: 460437]
12. Gelin BR, Karplus M. Proc. Natl. Acad. Sci. U.S.A 1975;72:2002. [PubMed: 1056008]
13. Ansari A, Berendzen J, Bowne SF, Frauenfelder H, Iben IE, Sauke TB, Shyamsunder E, Young RD. Proc. Natl. Acad. Sci. U.S.A 1985;82:5000. [PubMed: 3860839]
14. McCammon JA, Gelin BR, Karplus M, Wolynes PG. Nature 1976;262:325. [PubMed: 958384]
15. Frauenfelder H, McMahon BH, Austin RH, Chu K, Groves JT. Proc. Natl. Acad. Sci. U.S.A 2001;98:2370. [PubMed: 11226246]
16. Feynman, R. The Feynman lectures on physics. Addison-Wesley Publishing Company; Reading, MA: 1963.
17. Wittenberg JB, Wittenberg BA. Annu. Rev. Biophys. Biophys. Chem 1990;19:217. [PubMed: 2194476]
18. Shrivastava IH, Sansom MS. Biophys. J 2000;78:557. [PubMed: 10653771]
19. Berneche S, Roux B. Nature 2001;414:73. [PubMed: 11689945]
20. Roux B. Acc. Chem. Res 2002;35:366. [PubMed: 12069621]
21. Bahar I, Chennubhotla C, Tobi D. Curr. Opin. Struct. Biol 2007;17:633. [PubMed: 18024008]
22. Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D. Nature 2005;438:117. [PubMed: 16267559]
23. Henzler-Wildman K, Kern D. Nature 2007;450:964. [PubMed: 18075575]
24. Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, Pozharski E, Wilson MA, Petsko GA, Karplus M, Hubner CG, Kern D. Nature 2007;450:838. [PubMed: 18026086]
25. Kern D, Zuiderweg ER. Curr. Opin. Struct. Biol 2003;13:748. [PubMed: 14675554]
26. Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KF, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL. Science 2008;320:1471. [PubMed: 18556554]
27. Lena C, Changeux JP. Trends Neurosci 1993;16:181. [PubMed: 7685943]
28. Ming D, Wall ME. Proteins 2005;59:697. [PubMed: 15822100]
29. Taly A, Corringer PJ, Grutter T, Prado de CL, Karplus M, Changeux JP. Proc. Natl. Acad. Sci. U.S.A 2006;103:16965. [PubMed: 17077146]
30. Tobi D, Bahar I. Proc. Natl. Acad. Sci. U.S.A 2005;102:18908. [PubMed: 16354836]
31. Xu C, Tobi D, Bahar I. J. Mol. Biol 2003;333:153. [PubMed: 14516750]
32. Yang LW, Bahar I. Structure 2005;13:893. [PubMed: 15939021]
33. Tajkhorshid E, Nollert P, Jensen MO, Miercke LJ, O'Connell J, Stroud RM, Schulten K. Science 2002;296:525. [PubMed: 11964478]
34. Noskov SY, Berneche S, Roux B. Nature 2004;431:830. [PubMed: 15483608]
35. Humprey W, Dalket A, Schulten K. J. Mol. Graphics 1996;14:33.
36. Jolliffe, IT. Principal Component Analysis. Springer; New York: 2002.
37. Kitao A, Go N. Curr. Opin. Struct. Biol 1999;9:164. [PubMed: 10322205]
38. Bahar I, Rader AJ. Curr. Opin. Struct. Biol 2005;15:586. [PubMed: 16143512]

39. Cui, Q.; Bahar, IE. Normal Mode Analysis. Theory and Applications to Biological and Chemical Systems. Taylor & Francis Group; Boca Raton, FL: 2006.
40. Amadei A, Linssen AB, Berendsen HJ. Proteins 1993;17:412. [PubMed: 8108382]
41. Bahar I, Erman B, Haliloglu T, Jernigan RL. Biochemistry 1997;36:13512. [PubMed: 9354619]
42. Garcia AE, Harman JG. Protein Sci 1996;5:62. [PubMed: 8771197]
43. Romo TD, Clarage JB, Sorensen DC, Phillips GN Jr. Proteins 1995;22:311. [PubMed: 7479706]
44. Yang LW, Eyal E, Bahar I, Kitao A. Bioinformatics 2009;25:606. [PubMed: 19147661]
45. Chennubhotla C, Yang Z, Bahar I. Mol. BioSyst 2008;4:287. [PubMed: 18354781]
46. Henzler-Wildman KA, Lei M, Thai V, Kerns SJ, Karplus M, Kern D. Nature 2007;450:913. [PubMed: 18026087]
47. Swain JF, Gierasch LM. Curr. Opin. Struct. Biol 2006;16:102. [PubMed: 16423525]
48. Hawkins RJ, McLeish TC. Biophys. J 2006;91:2055. [PubMed: 16798805]
49. Popovych N, Sun S, Ebright RH, Kalodimos CG. Nat. Struct. Mol. Biol 2006;13:831. [PubMed: 16906160]
50. Brath U, Akke M, Yang D, Kay LE, Mulder FA. J. Am. Chem. Soc 2006;128:5718. [PubMed: 16637639]
51. Velyvis A, Yang YR, Schachman HK, Kay LE. Proc. Natl. Acad. Sci. U.S.A 2007;104:8815. [PubMed: 17502625]
52. Changeux JP, Edelstein SJ. Science 2005;308:1424. [PubMed: 15933191]
53. Monod J, Wyman J, Changeux JP. J. Mol. Biol 1965;12:88. [PubMed: 14343300]
54. Go N, Noguti T, Nishikawa T. Proc. Natl. Acad. Sci. U.S.A 1983;80:3696. [PubMed: 6574507]
55. Brooks B, Karplus M. Proc. Natl. Acad. Sci. U.S.A 1983;80:6571. [PubMed: 6579545]
56. Levitt M, Sander C, Stern PS. J. Mol. Biol 1985;181:423. [PubMed: 2580101]
57. Naik VM, Krimm S, Denton JB, Nemethy G, Scheraga HA. Int. J. Pept. Protein Res 1984;24:613. [PubMed: 6085071]
58. Tirion MM. Phys. Rev. Lett 1996;77:1905. [PubMed: 10063201]
59. Taylor, WR.; Aszodi, A. Protein Geometry, Classification, Topology and Symmetry: A Computational Analysis of Structure. Taylor & Francis Group; New York: 2005.
60. Lezon TR, Banavar JR, Lesk AM, Maritan A. Proteins 2006;63:273. [PubMed: 16470841]
61. Doruker P, Jernigan RL, Bahar I. J. Comput. Chem 2002;23:119. [PubMed: 11913377]
62. Lu M, Ma J. Biophys. J 2005;89:2395. [PubMed: 16055547]
63. Nicolay S, Sanejouand YH. Phys. Rev. Lett 2006;96:078104. [PubMed: 16606146]
64. Tama F, Brooks CL. Annu. Rev. Biophys. Biomol. Struct 2006;35:115. [PubMed: 16689630]
65. Zheng W, Brooks BR, Thirumalai D. Proc. Natl. Acad. Sci. U.S.A 2006;103:7664. [PubMed: 16682636]
66. Ma J. Structure 2005;13:373. [PubMed: 15766538]
67. Bahar I, Atilgan AR, Erman B. Fold. Des 1997;2:173. [PubMed: 9218955]
68. Haliloglu T, Bahar I, Erman B. Phys. Rev. Lett 1997;79:3090.
69. Eichinger BE. Macromolecules 1972;5:496.
70. Flory PJ. Proc. R. Soc. London, A 1976;351:351.
71. Pearson DS. Macromolecules 1977;10:696.
72. Kloczkowski A, Mark JE, Erman B. Macromolecules 1998;22:1423.
73. Erman, B.; Mark, JE. Structure and Properties of Rubberlike Networks. Oxford University Press; New York: 1997.
74. Hinsen K. Proteins 1998;33:417. [PubMed: 9829700]
75. Hinsen K, Thomas A, Field MJ. Proteins 1999;34:369. [PubMed: 10024023]
76. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Biophys. J 2001;80:505. [PubMed: 11159421]
77. Doruker P, Atilgan AR, Bahar I. Proteins 2000;40:512. [PubMed: 10861943]
78. Eyal E, Yang LW, Bahar I. Bioinformatics 2006;22:2619. [PubMed: 16928735]

79. Tama F, Sanejouand YH. Protein Eng 2001;14:1. [PubMed: 11287673]
80. Mouawad L, Perahia D. J. Mol. Biol 1996;258:393. [PubMed: 8627633]
81. Tama F, Valle M, Frank J, Brooks CL III. Proc. Natl. Acad. Sci. U.S.A 2003;100:9319. [PubMed: 12878726]
82. Wang Y, Rader AJ, Bahar I, Jernigan RL. J. Struct. Biol 2004;147:302. [PubMed: 15450299]
83. Essiz SG, Coalson RD. J. Chem. Phys 2007;127:104109. [PubMed: 17867739]
84. Haddadian EJ, Cheng MH, Coalson RD, Xu Y, Tang P. J. Phys. Chem. B 2008;112:13981. [PubMed: 18847252]
85. Isin B, Schulten K, Tajkhorshid E, Bahar I. Biophys. J 2008;95:789. [PubMed: 18390613]
86. Miller BT, Zheng W, Venable RM, Pastor RW, Brooks BR. J. Phys. Chem. B 2008;112:6274. [PubMed: 18311963]
87. Moritsugu K, Smith JC. Biophys. J 2007;93:3460. [PubMed: 17693469]
88. Sherwood P, Brooks BR, Sansom MS. Curr. Opin. Struct. Biol 2008;18:630. [PubMed: 18721882]
89. Hille, B. Ion channels of excitable membranes. Sinauer Associates; Sunderland, MA: 2001.
90. Staros JV. Acc. Chem. Res 1988;21:435.
91. Gouaux E, MacKinnon R. Science 2005;310:1461. [PubMed: 16322449]
92. Dobson PD, Kell DB. Nat. Rev. Drug Discov 2008;7:205. [PubMed: 18309312]
93. Stahlberg H, Fotiadis D, Scheuring S, Remigy H, Braun T, Mitsuoka K, Fujiyoshi Y, Engel A. FEBS Lett 2001;504:166. [PubMed: 11532449]
94. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z, Woolsey J. Nucleic Acids Res 2006;34:D668–D672. [PubMed: 16381955]
95. Chen X, Ji ZL, Chen YZ. Nucleic Acids Res 2002;30:412. [PubMed: 11752352]
96. Gunther S, Kuhn M, Dunkel M, Campillos M, Senger C, Petsalaki E, Ahmed J, Urdiales EG, Gewiess A, Jensen LJ, Schneider R, Skoblo R, Russell RB, Bourne PE, Bork P, Preissner R. Nucleic Acids Res 2008;36:D919–D922. [PubMed: 17942422]
97. Imming P, Sinning C, Meyer A. Nat. Rev. Drug Discov 2006;5:821. [PubMed: 17016423]
98. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A. Genome Res 2003;13:2129. [PubMed: 12952881]
99. Overington JP, Al-Lazikani B, Hopkins AL. Nat. Rev. Drug Discov 2006;5:993. [PubMed: 17139284]
100. Zheng CJ, Han LY, Yap CW, Ji ZL, Cao ZW, Chen YZ. Pharmacol. Rev 2006;58:259. [PubMed: 16714488]
101. Filmore D. Modern Drug Discovery 2004;17:24.
102. Congreve M, Murray CW, Blundell TL. Drug Discovery Today 2005;10:895. [PubMed: 15993809]
103. Ostermeier C, Michel H. Curr. Opin. Struct. Biol 1997;7:697. [PubMed: 9345629]
104. White SH. Protein Sci 2004;13:1948. [PubMed: 15215534]
105. Blow N. Nat. Methods 2008;5:203. [PubMed: 18235436]
106. Service RF. Science 2008;319:1612. [PubMed: 18356502]
107. Carpenter EP, Beis K, Cameron AD, Iwata S. Curr. Opin. Struct. Biol 2008;18:581. [PubMed: 18674618]
108. Arora A, Tamm LK. Curr. Opin. Struct. Biol 2001;11:540. [PubMed: 11785753]
109. Marassi FM, Opella SJ. Curr. Opin. Struct. Biol 1998;8:640. [PubMed: 9818270]
110. Creuzet F, McDermott A, Gebhard R, van der Hoef K, Spijker-Assink MB, Herzfeld J, Lugtenburg J, Levitt MH, Griffin RG. Science 1991;251:783. [PubMed: 1990439]
111. Ma D, Tillman TS, Tang P, Meirovitch E, Eckenhoff R, Carnini A, Xu Y. Proc. Natl. Acad. Sci. U.S.A 2008;105:16537. [PubMed: 18948596]
112. Marassi FM, Ramamoorthy A, Opella SJ. Proc. Natl. Acad. Sci. U.S.A 1997;94:8551. [PubMed: 9238014]
113. Marassi FM, Ma C, Gratkowski H, Straus SK, Strebel K, Oblatt-Montal M, Montal M, Opella SJ. Proc. Natl. Acad. Sci. U.S.A 1999;96:14336. [PubMed: 10588706]
114. Marassi FM, Opella SJ. J. Magn. Reson 2000;144:150. [PubMed: 10783285]
115. Marassi FM, Opella SJ. Protein Sci 2003;12:403. [PubMed: 12592011]

116. Opella SJ, Nevzorov A, Mesleb MF, Marassi FM. Biochem. Cell Biol 2002;80:597. [PubMed: 12440700]
117. Chang G, Spencer RH, Lee AT, Barclay MT, Rees DC. Science 1998;282:2220. [PubMed: 9856938]
118. Doyle DA, Morais CJ, Pfuetzner RA, Kuo A, Gulbis JM, Cohen SL, Chait BT, MacKinnon R. Science 1998;280:69. [PubMed: 9525859]
119. Jiang Y, Lee A, Chen J, Cadene M, Chait BT, MacKinnon R. Nature 2002;417:523. [PubMed: 12037560]
120. Jiang Y, Lee A, Chen J, Cadene M, Chait BT, MacKinnon R. Nature 2002;417:515. [PubMed: 12037559]
121. Yeh JI, Biemann HP, Prive GG, Pandit J, Koshland DE Jr. Kim SH. J. Mol. Biol 1996;262:186. [PubMed: 8831788]
122. Yeh JI. Acta Crystallogr., D: Biol. Crystallogr 2003;59:1408. [PubMed: 12876342]
123. Arrondo JL, Goni FM. Prog. Biophys. Mol. Biol 1999;72:367. [PubMed: 10605294]
124. Tusnady GE, Dosztanyi Z, Simon I. Nucleic Acids Res 2005;33:D275–D278. [PubMed: 15608195]
125. Raman P, Cherezov V, Caffrey M. Cell. Mol. Life Sci 2006;63:36. [PubMed: 16314922]
126. Sutherland GB. Adv. Protein Chem 1952;7:291. [PubMed: 14933255]
127. Fanucci GE, Cafiso DS. Curr. Opin. Struct. Biol 2006;16:644. [PubMed: 16949813]
128. Hubbell WL, Cafiso DS, Altenbach C. Nat. Struct. Biol 2000;7:735. [PubMed: 10966640]
129. Hubbell, WL.; Altenbach, C. Membrane Protein Structure: Experimental Approaches. Oxford University Press; New York: 1994. p. 224-248.
130. Lakomek NA, Walter KF, Fares C, Lange OF, de Groot BL, Grubmuller H, Bruschweiler R, Munk A, Becker S, Meiler J, Griesinger C. J. Biomol. NMR 2008;41:139. [PubMed: 18523727]
131. Mittermaier A, Kay LE. Science 2006;312:224. [PubMed: 16614210]
132. Sprangers R, Velyvis A, Kay LE. Nat. Methods 2007;4:697. [PubMed: 17762877]
133. Tugarinov V, Sprangers R, Kay LE. J. Am. Chem. Soc 2007;129:1743. [PubMed: 17249677]
134. Hubbell WL, Altenbach C, Hubbell CM, Khorana HG. Adv. Protein Chem 2003;63:243. [PubMed: 12629973]
135. Ishima R, Torchia DA. Nat. Struct. Biol 2000;7:740. [PubMed: 10966641]
136. Kay LE. J. Magn. Reson 2005;173:193. [PubMed: 15780912]
137. Labeikovsky W, Eisenmesser EZ, Bosco DA, Kern D. J. Mol. Biol 2007;367:1370. [PubMed: 17316687]
138. Akke M, Skelton NJ, Kordel J, Palmer AG III, Chazin WJ. Biochemistry 1993;32:9832. [PubMed: 8373781]
139. Akke M, Fiala R, Jiang F, Patel D, Palmer AG III. RNA 1997;3:702. [PubMed: 9214654]
140. Akke M, Liu J, Cavanagh J, Erickson HP, Palmer AG III. Nat. Struct. Biol 1998;5:55. [PubMed: 9437430]
141. Tugarinov V, Kay LE. Biochemistry 2005;44:15970. [PubMed: 16331956]
142. Tugarinov V, Kay LE. J. Am. Chem. Soc 2006;128:12484. [PubMed: 16984199]
143. Sprangers R, Kay LE. Nature 2007;445:618. [PubMed: 17237764]
144. Thomas A, Field MJ, Mouawad L, Perahia D. J. Mol. Biol 1996;257:1070. [PubMed: 8632469]
145. Velyvis A, Yang YR, Schachman HK, Kay LE. Proc. Natl. Acad. Sci. U.S.A 2007;104:8815. [PubMed: 17502625]
146. Tolman JR, Flanagan JM, Kennedy MA, Prestegard JH. Nat. Struct. Biol 1997;4:292. [PubMed: 9095197]
147. Hubbell WL, Mchaourab HS, Altenbach C, Lietzow MA. Structure 1996;4:779. [PubMed: 8805569]
148. Altenbach C, Marti T, Khorana HG, Hubbell WL. Science 1990;248:1088. [PubMed: 2160734]
149. Steinhoff HJ, Mollaaghababa R, Altenbach C, Hideg K, Krebs M, Khorana HG, Hubbell WL. Science 1994;266:105. [PubMed: 7939627]
150. Altenbach C, Yang K, Farrens DL, Farahbakhsh ZT, Khorana HG, Hubbell WL. Biochemistry 1996;35:12470. [PubMed: 8823182]

151. Altenbach C, Klein-Seetharaman J, Cai K, Khorana HG, Hubbell WL. Biochemistry 2001;40:15493. [PubMed: 11747424]
152. Altenbach C, Cai K, Klein-Seetharaman J, Khorana HG, Hubbell WL. Biochemistry 2001;40:15483. [PubMed: 11747423]
153. Lewis JW, van Kuijk FJ, Thorgeirsson TE, Kliger DS. Biochemistry 1991;30:11372. [PubMed: 1742277]
154. Resek JF, Farahbakhsh ZT, Hubbell WL, Khorana HG. Biochemistry 1993;32:12025. [PubMed: 8218279]
155. Steinhoff HJ, Schwemer J. J. Photochem. Photobiol. B 1996;35:1. [PubMed: 8823930]
156. Liu YS, Sompornpisut P, Perozo E. Nat. Struct. Biol 2001;8:883. [PubMed: 11573095]
157. Perozo E, Cortes DM, Cuello LG. Science 1999;285:73. [PubMed: 10390363]
158. Perozo E, Cuello LG, Cortes DM, Liu YS, Sompornpisut P. Novartis Found. Symp 2002;245:146. [PubMed: 12027005]
159. Sompornpisut P, Liu YS, Perozo E. Biophys. J 2001;81:2530. [PubMed: 11606268]
160. Perozo E, Cortes DM, Sompornpisut P, Kloda A, Martinac B. Nature 2002;418:942. [PubMed: 12198539]
161. Altenbach C, Kusnetzow AK, Ernst OP, Hofmann KP, Hubbell WL. Proc. Natl. Acad. Sci. U.S.A 2008;105:7439. [PubMed: 18490656]
162. Doniach S, Eastman P. Curr. Opin. Struct. Biol 1999;9:157. [PubMed: 10322213]
163. Khalili-Araghi F, Gumbart J, Wen PC, Sotomayor M, Tajkhorshid E, Schulten K. Curr. Opin. Struct. Biol 2009;19:128. [PubMed: 19345092]
164. Tieleman DP, Biggin PC, Smith GR, Sansom MS. Q. Rev. Biophys 2001;34:473. [PubMed: 11852594]
165. Ash WL, Zlomislic MR, Oloo EO, Tieleman DP. Biochim. Biophys. Acta 2004;1666:158. [PubMed: 15519314]
166. Tieleman DP. Clin. Exp. Pharmacol. Physiol 2006;33:893. [PubMed: 17002665]
167. Stone JE, Phillips JC, Freddolino PL, Hardy DJ, Trabuco LG, Schulten K. J. Comput. Chem 2007;28:2618. [PubMed: 17894371]
168. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. J. Comput. Chem 2005;26:1781. [PubMed: 16222654]
169. Hub JS, de Groot BL. Proc. Natl. Acad. Sci. U.S.A 2008;105:1198. [PubMed: 18202181]
170. Wang Y, Schulten K, Tajkhorshid E. Structure 2005;13:1107. [PubMed: 16084383]
171. Tornroth-Horsefield S, Wang Y, Hedfalk K, Johanson U, Karlsson M, Tajkhorshid E, Neutze R, Kjellbom P. Nature 2006;439:688. [PubMed: 16340961]
172. de Groot BL, Grubmuller H. Science 2001;294:2353. [PubMed: 11743202]
173. Allen TW, Bastug T, Kuyucak S, Chung SH. Biophys. J 2003;84:2159. [PubMed: 12668425]
174. Allen TW, Andersen OS, Roux B. Biophys. J 2006;90:3447. [PubMed: 16500984]
175. Bastug T, Patra SM, Kuyucak S. Chem. Phys. Lipids 2006;141:197. [PubMed: 16600199]
176. Roux B, MacKinnon R. Science 1999;285:100. [PubMed: 10390357]
177. Shrivastava IH, Sansom MS. Eur. Biophys. J 2002;31:207. [PubMed: 12029333]
178. Shrivastava IH, Tieleman DP, Biggin PC, Sansom MS. Biophys. J 2002;83:633. [PubMed: 12124253]
179. Allen TW, Kuyucak S, Chung SH. Biophys. J 1999;77:2502. [PubMed: 10545352]
180. Grossfield A, Feller SE, Pitman MC. Proteins 2007;67:31. [PubMed: 17243153]
181. Roux B, Allen T, Berneche S, Im W. Q. Rev. Biophys 2004;37:15. [PubMed: 17390604]
182. Goldstein, H. Classical Mechanics. Addison-Wesley; Cambridge: 1953.
183. Song G, Jernigan RL. Proteins 2006;63:197. [PubMed: 16447281]
184. Yang L, Song G, Jernigan RL. Biophys. J 2007;93:920. [PubMed: 17483178]
185. Chu JW, Voth GA. Biophys. J 2006;90:1572. [PubMed: 16361345]
186. Jang Y, Jeong JI, Kim MK. Nucleic Acids Res 2006;34:W57–W62. [PubMed: 16845072]

187. Kim MK, Li W, Shapiro BA, Chirikjian GS. J. Biomol. Struct. Dyn 2003;21:395. [PubMed: 14616035]
188. Liu X, Xu Y, Li H, Wang X, Jiang H, Barrantes FJ. PLoS Comput. Biol 2008;4:e19. [PubMed: 18225945]
189. Miloshevsky GV, Jordan PC. Structure 2007;15:1654. [PubMed: 18073114]
190. Yang L, Song G, Carriquiry A, Jernigan RL. Structure 2008;16:321. [PubMed: 18275822]
191. Nichols J, Taylor H, Schmidt P, Simons J. J. Chem. Phys 1990;92:340.
192. Miloshevsky GV, Jordan PC. Structure 2006;14:1241. [PubMed: 16905098]
193. Lu M, Ma J. Proc. Natl. Acad. Sci. U.S.A 2008;105:15358. [PubMed: 18832168]
194. Kundu S, Melton JS, Sorensen DC, Phillips GN Jr. Biophys. J 2002;83:723. [PubMed: 12124259]
195. Sen, TZ.; Jernigan, RL. Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems. Chapman & Hall/CRC; Boca Raton, FL: 2006. p. 171-186.
196. Kondrashov DA, Cui Q, Phillips GN Jr. Biophys. J 2006;91:2760. [PubMed: 16891367]
197. Hamacher K, McCammon JA. J. Chem. Theory Comput 2006;2:873.
198. Lyman E, Pfaendtner J, Voth GA. Biophys. J 2008;95:4183. [PubMed: 18658214]
199. Chennubhotla C, Bahar I. Mol. Syst. Biol 2006;2:36. [PubMed: 16820777]
200. Franklin J, Koehl P, Doniach S, Delarue M. Nucleic Acids Res 2007;35:W477–W482. [PubMed: 17545201]
201. Lindahl E, Azuara C, Koehl P, Delarue M. Nucleic Acids Res 2006;34:W52–W56. [PubMed: 16845062]
202. Suhre K, Sanejouand YH. Nucleic Acids Res 2004;32:W610–W614. [PubMed: 15215461]
203. Wako H, Kato M, Endo S. Bioinformatics 2004;20:2035. [PubMed: 15059828]
204. Yang LW, Liu X, Jursa CJ, Holliman M, Rader A, Karimi HA, Bahar I. Bioinformatics 2005;21:2978. [PubMed: 15860562]
205. Yang LW, Rader AJ, Liu X, Jursa CJ, Chen SC, Karimi HA, Bahar I. Nucleic Acids Res 2006;34:W24–W31. [PubMed: 16845002]
206. Bahar I, Atilgan AR, Demirel MC, Erman B. Phys. Rev. Lett 1998;80:2733.
207. Chennubhotla C, Rader AJ, Yang LW, Bahar I. Phys. Biol 2005;2:S173–S180. [PubMed: 16280623]
208. Rader, AJ.; Chennubhotla, C.; Yang, LW.; Bahar, I. Normal Mode Analysis. Theory and Applications to Biological and Chemical Systems. Taylor & Francis Group; New York: 2006. p. 41-64.Chapter 3
209. Yang LW, Eyal E, Chennubhotla C, Jee J, Gronenborn AM, Bahar I. Structure 2007;15:741. [PubMed: 17562320]
210. Bahar I, Wallqvist A, Covell DG, Jernigan RL. Biochemistry 1998;37:1067. [PubMed: 9454598]
211. Haliloglu T, Bahar I. Proteins 1999;37:654. [PubMed: 10651280]
212. Temiz NA, Meirovitch E, Bahar I. Proteins 2004;57:468. [PubMed: 15382240]
213. Zhuravleva A, Korzhnev DM, Nolde SB, Kay LE, Arseniev AS, Billeter M, Orekhov VY. J. Mol. Biol 2007;367:1079. [PubMed: 17306298]
214. Su JG, Li CH, Hao R, Chen WZ, Wang CX. Biophys. J 2008;94:4586. [PubMed: 18310247]
215. Rader AJ, Bahar I. Polymer 2004;45:659.
216. Rader AJ, Anderson G, Isin B, Khorana HG, Bahar I, Klein-Seetharaman J. Proc. Natl. Acad. Sci. U.S.A 2004;101:7246. [PubMed: 15123809]
217. Maguid S, Fernandez-Alberti S, Ferrelli L, Echave J. Biophys. J 2005;89:3. [PubMed: 15749782]
218. Shrivastava IH, Bahar I. Biophys. J 2006;90:3929. [PubMed: 16533848]
219. Marques O, Sanejouand YH. Proteins 1995;23:557. [PubMed: 8749851]
220. Bruschweiler R. J. Chem. Phys 1995;102:3396.
221. Ayton GS, Noid WG, Voth GA. Curr. Opin. Struct. Biol 2007;17:192. [PubMed: 17383173]
222. Kurkcuoglu O, Jernigan RL, Doruker P. Biochemistry 2006;45:1173. [PubMed: 16430213]
223. Noid WG, Chu JW, Ayton GS, Krishna V, Izvekov S, Voth GA, Das A, Andersen HC. J. Chem. Phys 2008;128:244114. [PubMed: 18601324]
224. van Vlijmen HW, Karplus M. J. Chem. Phys 2001;115:698.

225. van Vlijmen HW, Karplus M. J. Mol. Biol 2005;350:528. [PubMed: 15922356]
226. Yang Z, Bahar I, Widom M. Biophys. J 2009;96:4438. [PubMed: 19486668]
227. Kim MK, Jernigan RL, Chirikjian GS. J. Struct. Biol 2003;143:107. [PubMed: 12972347]
228. Tama F, Gadea FX, Marques O, Sanejouand YH. Proteins 2000;41:1. [PubMed: 10944387]
229. Li G, Cui Q. Biophys. J 2002;83:2457. [PubMed: 12414680]
230. Durand P, Trinquier G, Sanejouand YH. Biopolymers 1994;34:759.
231. Li G, Cui Q. Biophys. J 2004;86:743. [PubMed: 14747312]
232. Tama F, Brooks CL III. J. Mol. Biol 2002;318:733. [PubMed: 12054819]
233. Tama F, Brooks CL III. J. Mol. Biol 2005;345:299. [PubMed: 15571723]
234. Hinsen K, Kneller GR. Proteins 2008;70:1235. [PubMed: 17853448]
235. Miller BT, Zheng W, Venable RM, Pastor RW, Brooks BR. J. Phys. Chem. B 2008;112:6274. [PubMed: 18311963]
236. Moritsugu K, Smith JC. J. Phys. Chem. B 2005;109:12182. [PubMed: 16852503]
237. Zheng W, Brooks BR. Biophys. J 2005;89:167. [PubMed: 15879477]
238. Ming D, Wall ME. Phys. Rev. Lett 2005;95:198103. [PubMed: 16384030]
239. Ming D, Wall ME. J. Mol. Biol 2006;358:213. [PubMed: 16513135]
240. Sonne J, Kandt C, Peters GH, Hansen FY, Jensen MO, Tieleman DP. Biophys. J 2007;92:2727. [PubMed: 17208973]
241. Lamm G, Szabo A. J. Chem. Phys 1986;85:7334.
242. Moritsugu K, Smith JC. J. Phys. Chem. B 2006;110:5807. [PubMed: 16539528]
243. Kneller GR. Chem. Phys 2000;261:1.
244. Hinsen K, Petrescu AJ, Dellerue S, Bellissent-Funel MC, Kneller GR. J. Chem. Phys 2000;261:25.
245. Roux B, Karplus M. Biophys. J 1988;53:297. [PubMed: 2450595]
246. Lu HP. Acc. Chem. Res 2005;38:557. [PubMed: 16028890]
247. Harms GS, Orr G, Montal M, Thrall BD, Colson SD, Lu HP. Biophys. J 2003;85:1826. [PubMed: 12944296]
248. Thompson JD, Higgins DG, Gibson TJ. Nucleic Acids Res 1994;22:4673. [PubMed: 7984417]
249. Shen Y, Kong Y, Ma J. Proc. Natl. Acad. Sci. U.S.A 2002;99:1949. [PubMed: 11842204]
250. Shimizu H, Iwamoto M, Konno T, Nihei A, Sasaki YC, Oiki S. Cell 2008;132:67. [PubMed: 18191221]
251. Alam A, Jiang Y. Nat. Struct. Mol. Biol 2009;16:35. [PubMed: 19098915]
252. Shi N, Ye S, Alam A, Chen L, Jiang Y. Nature 2006;440:570. [PubMed: 16467789]
253. Haliloglu T, Ben-Tal N. PLoS Comput. Biol 2008;4:e1000164. [PubMed: 18769593]
254. Yifrach O, MacKinnon R. Cell 2002;111:231. [PubMed: 12408867]
255. Jan LY, Jan YN. Nature 1994;371:119. [PubMed: 8072541]
256. Anishkin A, Chiang CS, Sukharev S. J. Gen. Physiol 2005;125:155. [PubMed: 15684093]
257. Hamill OP, Martinac B. Physiol. Rev 2001;81:685. [PubMed: 11274342]
258. Sukharev S, Durell SR, Guy HR. Biophys. J 2001;81:917. [PubMed: 11463635]
259. Poolman B, Blount P, Folgering JH, Friesen RH, Moe PC, van der Heide T. Mol. Microbiol 2002;44:889. [PubMed: 12010487]
260. Cruickshank CC, Minchin RF, Le Dain AC, Martinac B. Biophys. J 1997;73:1925. [PubMed: 9336188]
261. Valadie H, Lacapere JJ, Sanejouand YH, Etchebest C. J. Mol. Biol 2003;332:657. [PubMed: 12963374]
262. Czajkowski C. Nature 2005;438:167. [PubMed: 16281019]
263. Sine SM, Engel AG. Nature 2006;440:448. [PubMed: 16554804]
264. Miyazawa A, Fujiyoshi Y, Unwin N. Nature 2003;423:949. [PubMed: 12827192]
265. Unwin N. J. Mol. Biol 2005;346:967. [PubMed: 15701510]
266. Brejc K, van Dijk WJ, Klaassen RV, Schuurmans M, van der Oost J, Smit AB, Sixma TK. Nature 2001;411:269. [PubMed: 11357122]

267. Changeux JP, Edelstein SJ. Neuron 1998;21:959. [PubMed: 9856454]
268. Hung A, Tai K, Sansom MS. Biophys. J 2005;88:3321. [PubMed: 15722430]
269. Liu X, Xu Y, Li H, Wang X, Jiang H, Barrantes FJ. PLoS Comput. Biol 2008;4:e19. [PubMed: 18225945]
270. Szarecka A, Xu Y, Tang P. Proteins 2007;68:948. [PubMed: 17546671]
271. Taly A, Delarue M, Grutter T, Nilges M, Le NN, Corringer PJ, Changeux JP. Biophys. J 2005;88:3954. [PubMed: 15805177]
272. Cheng X, Lu B, Grant B, Law RJ, McCammon JA. J. Mol. Biol 2006;355:310. [PubMed: 16307758]
273. Smart OS, Neduvelil JG, Wang X, Wallace BA, Sansom MS. J. Mol. Graph 1996;14:354. [PubMed: 9195488]
274. Corry B. Biophys. J 2006;90:799. [PubMed: 16284265]
275. Unwin N. Nature 1995;373:37. [PubMed: 7800037]
276. Cymes GD, Ni Y, Grosman C. Nature 2005;438:975. [PubMed: 16355215]
277. Cymes GD, Grosman C. Nat. Struct. Mol. Biol 2008;15:389. [PubMed: 18376414]
278. Bouzat C, Gumilar F, Spitzmaul G, Wang HL, Rayes D, Hansen SB, Taylor P, Sine SM. Nature 2004;430:896. [PubMed: 15318223]
279. Lee WY, Sine SM. Nature 2005;438:243. [PubMed: 16281039]
280. Auerbach A. Proc. Natl. Acad. Sci. U.S.A 2005;102:1408. [PubMed: 15665102]
281. Purohit P, Mitra A, Auerbach A. Nature 2007;446:930. [PubMed: 17443187]
282. Mukhtasimova N, Lee WY, Wang HL, Sine SM. Nature 2009;459:451. [PubMed: 19339970]
283. Hilf RJ, Dutzler R. Nature 2008;452:375. [PubMed: 18322461]
284. Bocquet N, Nury H, Baaden M, Le PC, Changeux JP, Delarue M, Corringer PJ. Nature 2009;457:111. [PubMed: 18987633]
285. Hilf RJ, Dutzler R. Nature 2009;457:115. [PubMed: 18987630]
286. Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB. Mol. Pharmacol 2003;63:1256. [PubMed: 12761335]
287. Kobilka BK. Biochim. Biophys. Acta 2007;1768:794. [PubMed: 17188232]
288. Oldham WM, Hamm HE. Nat. Rev. Mol. Cell Biol 2008;9:60. [PubMed: 18043707]
289. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le TI, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M. Science 2000;289:739. [PubMed: 10926528]
290. Hanson MA, Stevens RC. Structure 2009;17:8. [PubMed: 19141277]
291. Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, Choe HW, Hofmann KP, Ernst OP. Nature 2008;455:497. [PubMed: 18818650]
292. Park JH, Scheerer P, Hofmann KP, Choe HW, Ernst OP. Nature 2008;454:183. [PubMed: 18563085]
293. Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC. Science 2008;322:1211. [PubMed: 18832607]
294. Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF. Nature 2008;454:486. [PubMed: 18594507]
295. Murakami M, Kouyama T. Nature 2008;453:363. [PubMed: 18480818]
296. Kusnetzow AK, Altenbach C, Hubbell WL. Biochemistry 2006;45:5538. [PubMed: 16634635]
297. Farrens DL, Altenbach C, Yang K, Hubbell WL, Khorana HG. Science 1996;274:768. [PubMed: 8864113]
298. Farahbakhsh ZT, Altenbach C, Hubbell WL. Photochem. Photobiol 1992;56:1019. [PubMed: 1492127]
299. Bhattacharya S, Hall SE, Vaidehi N. J. Mol. Biol 2008;382:539. [PubMed: 18638482]
300. Isin B, Rader AJ, Dhiman HK, Klein-Seetharaman J, Bahar I. Proteins 2006;65:970. [PubMed: 17009319]
301. Shacham S, Marantz Y, Bar-Haim S, Kalid O, Warshaviak D, Avisar N, Inbal B, Heifetz A, Fichman M, Topf M, Naor Z, Noiman S, Becker OM. Proteins 2004;57:51. [PubMed: 15326594]
302. Lemaitre V, Yeagle P, Watts A. Biochemistry 2005;44:12667. [PubMed: 16171381]
303. Huber T, Botelho AV, Beyer K, Brown MF. Biophys. J 2004;86:2078. [PubMed: 15041649]

304. Crozier PS, Stevens MJ, Forrest LR, Woolf TB. J. Mol. Biol 2003;333:493. [PubMed: 14556740]
305. Rohrig UF, Guidoni L, Rothlisberger U. Biochemistry 2002;41:10799. [PubMed: 12196019]
306. Saam J, Tajkhorshid E, Hayashi S, Schulten K. Biophys. J 2002;83:3097. [PubMed: 12496081]
307. Okada T, Fujiyoshi Y, Silow M, Navarro J, Landau EM, Shichida Y. Proc. Natl. Acad. Sci. U.S.A 2002;99:5982. [PubMed: 11972040]
308. Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. J. Mol. Biol 2004;342:571. [PubMed: 15327956]
309. Salom D, Lodowski DT, Stenkamp RE, Le TI, Golczak M, Jastrzebska B, Harris T, Ballesteros JA, Palczewski K. Proc. Natl. Acad. Sci. U.S.A 2006;103:16123. [PubMed: 17060607]
310. Stenkamp RE. Acta Crystallogr., D: Biol. Crystallogr 2008;D64:902. [PubMed: 18645239]
311. Teller DC, Okada T, Behnke CA, Palczewski K, Stenkamp RE. Biochemistry 2001;40:7761. [PubMed: 11425302]
312. Nakamichi H, Okada T. Proc. Natl. Acad. Sci. U.S.A 2006;103:12729. [PubMed: 16908857]
313. Nakamichi H, Okada T. Angew. Chem., Int. Ed 2006;45:4270.
314. Nakamichi H, Buss V, Okada T. Biophys. J 2007;92:L106–L108. [PubMed: 17449675]
315. Standfuss J, Xie G, Edwards PC, Burghammer M, Oprian DD, Schertler GF. J. Mol. Biol 2007;372:1179. [PubMed: 17825322]
316. Zerangue N, Kavanaugh MP. Nature 1996;383:634. [PubMed: 8857541]
317. Tanaka K, Watase K, Manabe T, Yamada K, Watanabe M, Takahashi K, Iwama H, Nishikawa T, Ichihara N, Kikuchi T, Okuyama S, Kawashima N, Hori S, Takimoto M, Wada K. Science 1997;276:1699. [PubMed: 9180080]
318. Amara SG, Fontana AC. Neurochem. Int 2002;41:313. [PubMed: 12176072]
319. Saier MH Jr. Tran CV, Barabote RD. Nucleic Acids Res 2006;34:D181–D186. [PubMed: 16381841]
320. Slotboom DJ, Konings WN, Lolkema JS. Microbiol. Mol. Biol. Rev 1999;63:293. [PubMed: 10357852]
321. Amara SG. Nat. Struct. Mol. Biol 2007;14:792. [PubMed: 17823609]
322. Kanner BI, Borre L. Biochim. Biophys. Acta 2002;1555:92. [PubMed: 12206897]
323. Boudker O, Ryan RM, Yernool D, Shimamoto K, Gouaux E. Nature 2007;445:387. [PubMed: 17230192]
324. Yernool D, Boudker O, Jin Y, Gouaux E. Nature 2004;431:811. [PubMed: 15483603]
325. Torres GE, Amara SG. Curr. Opin. Neurobiol 2007;17:304. [PubMed: 17509873]
326. Shrivastava IH, Jiang J, Amara SG, Bahar I. J. Biol. Chem 2008;283:28680. [PubMed: 18678877]
327. Larsson HP, Tzingounis AV, Koch HP, Kavanaugh MP. Proc. Natl. Acad. Sci. U.S.A 2004;101:3951. [PubMed: 15001707]
328. Gu Y, Shrivastava IH, Amara SG, Bahar I. Proc. Natl. Acad. Sci. U.S.A 2009;106:2589. [PubMed: 19202063]
329. Mitrovic AD, Plesko F, Vandenberg RJ. J. Biol. Chem 2001;276:26071. [PubMed: 11352900]
330. Higgins CF. Annu. Rev. Cell Biol 1992;8:67. [PubMed: 1282354]
331. Higgins CF. Res. Microbiol 2001;152:205. [PubMed: 11421269]
332. Locher KP, Lee AT, Rees DC. Science 2002;296:1091. [PubMed: 12004122]
333. Locher KP. Curr. Opin. Struct. Biol 2004;14:426. [PubMed: 15313236]
334. Locher KP, Borths E. FEBS Lett 2004;564:264. [PubMed: 15111107]
335. Oloo EO, Tieleman DP. J. Biol. Chem 2004;279:45013. [PubMed: 15308647]
336. Oloo EO, Fung EY, Tieleman DP. J. Biol. Chem 2006;281:28397. [PubMed: 16877382]
337. Ivetac A, Campbell JD, Sansom MS. Biochemistry 2007;46:2767. [PubMed: 17302441]
338. Weng J, Ma J, Fan K, Wang W. Biophys. J 2008;94:612. [PubMed: 17951296]
339. Rader AJ, Vlad DH, Bahar I. Structure 2005;13:413. [PubMed: 15766543]
340. Kim MK, Jernigan RL, Chirikjian GS. Biophys. J 2002;83:1620. [PubMed: 12202386]
341. Kim MK, Jernigan RL, Chirikjian GS. Biophys. J 2005;89:43. [PubMed: 15833998]
342. Zheng W, Brooks BR. Biophys. J 2005;88:3109. [PubMed: 15722427]

343. Zheng W, Brooks BR. Biophys. J 2006;90:4327. [PubMed: 16565046]
344. Maragakis P, Karplus M. J. Mol. Biol 2005;352:807. [PubMed: 16139299]
345. Miyashita O, Onuchic JN, Wolynes PG. Proc. Natl. Acad. Sci. U.S.A 2003;100:12570. [PubMed: 14566052]
346. Yang Z, Májek P, Bahar I. PLoS Comput. Biol 2009;5:e1000360. [PubMed: 19381265]
347. Eyal E, Chennubhotla C, Yang LW, Bahar I. Bioinformatics 2007;23:i175–i184. [PubMed: 17646294]
348. Hinsen K. Bioinformatics 2008;24:521. [PubMed: 18089618]
349. Song G, Jernigan RL. J. Mol. Biol 2007;369:880. [PubMed: 17451743]
350. Liu L, Koharudin LMI, Gronenborn AM, Bahar I. Proteins. in press.
351. Bakan A, Bahar I. Proc. Natl. Acad. Sci. U.S.A 2009;106:14349. [PubMed: 19706521]
352. Li H, Helling R, Tang C, Wingreen N. Science 1996;273:666. [PubMed: 8662562]
353. Haider S, Grottesi A, Hall BA, Ashcroft FM, Sansom MS. Biophys. J 2005;88:3310. [PubMed: 15749783]
354. de Vries AH, Yefimov S, Mark AE, Marrink SJ. Proc. Natl. Acad. Sci. U.S.A 2005;102:5392. [PubMed: 15809443]
355. Faller R, Marrink SJ. Langmuir 2004;20:7686. [PubMed: 15323520]
356. Marrink SJ, de Vries AH, Mark AE. J. Phys. Chem. B 2004;108:750.
357. Bond PJ, Holyoake J, Ivetac A, Khalid S, Sansom MS. J. Struct. Biol 2007;157:593. [PubMed: 17116404]
358. Scott KA, Bond PJ, Ivetac A, Chetwynd AP, Khalid S, Sansom MS. Structure 2008;16:621. [PubMed: 18400182]
359. Lee AG. Mol. BioSyst 2005;1:203. [PubMed: 16880984]
360. Powl AM, East JM, Lee AG. Biochemistry 2003;42:14306. [PubMed: 14640699]
361. Powl AM, Wright JN, East JM, Lee AG. Biochemistry 2005;44:5713. [PubMed: 15823029]
362. Brooijmans N, Kuntz ID. Annu. Rev. Biophys. Biomol. Struct 2003;32:335. [PubMed: 12574069]
363. Shoichet BK, McGovern SL, Wei B, Irwin JJ. Curr. Opin. Chem. Biol 2002;6:439. [PubMed: 12133718]
364. Becker OM, Marantz Y, Shacham S, Inbal B, Heifetz A, Kalid O, Bar-Haim S, Warshaviak D, Fichman M, Noiman S. Proc. Natl. Acad. Sci. U.S.A 2004;101:11304. [PubMed: 15277683]
365. Varady J, Wu X, Fang X, Min J, Hu Z, Levant B, Wang S. J. Med. Chem 2003;46:4377. [PubMed: 14521403]
366. Kenakin T. Trends Pharmacol. Sci 2003;24:346. [PubMed: 12871667]
367. Ghanouni P, Steenhuis JJ, Farrens DL, Kobilka BK. Proc. Natl. Acad. Sci. U.S.A 2001;98:5997. [PubMed: 11353823]
368. Peleg G, Ghanouni P, Kobilka BK, Zare RN. Proc. Natl. Acad. Sci. U.S.A 2001;98:8469. [PubMed: 11438704]
369. Swaminath G, Xiang Y, Lee TW, Steenhuis J, Parnot C, Kobilka BK. J. Biol. Chem 2004;279:686. [PubMed: 14559905]
370. Bottegoni G, Kufareva I, Totrov M, Abagyan R. J. Med. Chem 2009;52:397. [PubMed: 19090659]
371. Totrov M, Abagyan R. Curr. Opin. Struct. Biol 2008;18:178. [PubMed: 18302984]
372. Cavasotto CN, Orry AJ, Abagyan RA. Proteins 2003;51:423. [PubMed: 12696053]
373. Cavasotto CN, Orry AJ, Murgolo NJ, Czarniecki MF, Kocsi SA, Hawes BE, O'Neill KA, Hine H, Burton MS, Voigt JH, Abagyan RA, Bayne ML, Monsma FJ Jr. J. Med. Chem 2008;51:581. [PubMed: 18198821]
374. Reynolds KA, Katritch V, Abagyan R. J. Comput. Aided Mol. Des 2009;23:273. [PubMed: 19148767]
375. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC. Science 2007;318:1258. [PubMed: 17962520]
376. Carlson HA, McCammon JA. Mol. Pharmacol 2000;57:213. [PubMed: 10648630]
377. Cavasotto CN, Abagyan RA. J. Mol. Biol 2004;337:209. [PubMed: 15001363]

378. Ferrari AM, Wei BQ, Costantino L, Shoichet BK. J. Med. Chem 2004;47:5076. [PubMed: 15456251]
379. May A, Zacharias M. Biochim. Biophys. Acta 2005;1754:225. [PubMed: 16214429]
380. Ota N, Agard DA. J. Mol. Biol 2001;314:607. [PubMed: 11846570]
381. Wei BQ, Weaver LH, Ferrari AM, Matthews BW, Shoichet BK. J. Mol. Biol 2004;337:1161. [PubMed: 15046985]
382. Cavasotto CN, Kovacs JA, Abagyan RA. J. Am. Chem. Soc 2005;127:9632. [PubMed: 15984891]
383. Floquet N, Marechal JD, Badet-Denisot MA, Robert CH, Dauchez M, Perahia D. FEBS Lett 2006;580:5130. [PubMed: 16962102]
384. May A, Zacharias M. J. Med. Chem 2008;51:3499. [PubMed: 18517186]
385. Zacharias M, Sklenar H. J. Comput. Chem 1999;20:287.
386. Floquet N, Dedieu S, Martiny L, Dauchez M, Perahia D. Arch. Biochem. Biophys 2008;478:103. [PubMed: 18675774]
387. Floquet N, Durand P, Maigret B, Badet B, Badet-Denisot MA, Perahia D. J. Mol. Biol 2009;385:653. [PubMed: 18976669]
388. Stumpff-Kane AW, Maksimiak K, Lee MS, Feig M. Proteins 2008;70:1345. [PubMed: 17876825]
389. Delarue M, Dumas P. Proc. Natl. Acad. Sci. U.S.A 2004;101:6957. [PubMed: 15096585]
390. Gorba C, Miyashita O, Tama F. Biophys. J 2008;94:1589. [PubMed: 17993489]
391. Hinsen K, Reuter N, Navaza J, Stokes DL, Lacapere JJ. Biophys. J 2005;88:818. [PubMed: 15542555]
392. Schroder GF, Brunger AT, Levitt M. Structure 2007;15:1630. [PubMed: 18073112]
393. Tama F, Miyashita O, Brooks CL III. J. Struct. Biol 2004;147:315. [PubMed: 15450300]
394. Tama F, Miyashita O, Brooks CL III. J. Mol. Biol 2004;337:985. [PubMed: 15033365]
395. Marechal JD, Perahia D. Eur. Biophys. J 2008;37:1157. [PubMed: 18286277]
396. Suhre K, Navaza J, Sanejouand YH. Acta Crystallogr., D: Biol. Crystallogr 2006;62:1098. [PubMed: 16929111]
397. Ming D, Kong Y, Lambert MA, Huang Z, Ma J. Proc. Natl. Acad. Sci. U.S.A 2002;99:8620. [PubMed: 12084922]
398. Tama F, Wriggers W, Brooks CL III. J. Mol. Biol 2002;321:297. [PubMed: 12144786]
399. Brink J, Ludtke SJ, Kong Y, Wakil SJ, Ma J, Chiu W. Structure 2004;12:185. [PubMed: 14962379]
400. Balog E, Smith JC, Perahia D. Phys. Chem. Chem. Phys 2006;8:5543. [PubMed: 17136269]
401. Best RB, Chen YG, Hummer G. Structure 2005;13:1755. [PubMed: 16338404]
402. Chu JW, Voth GA. Biophys. J 2007;93:3860. [PubMed: 17704151]
403. Yang H, Yu Y, Li WG, Yu F, Cao H, Xu TL, Jiang H. PLoS Biol 2009;7:e1000151. [PubMed: 19597538]

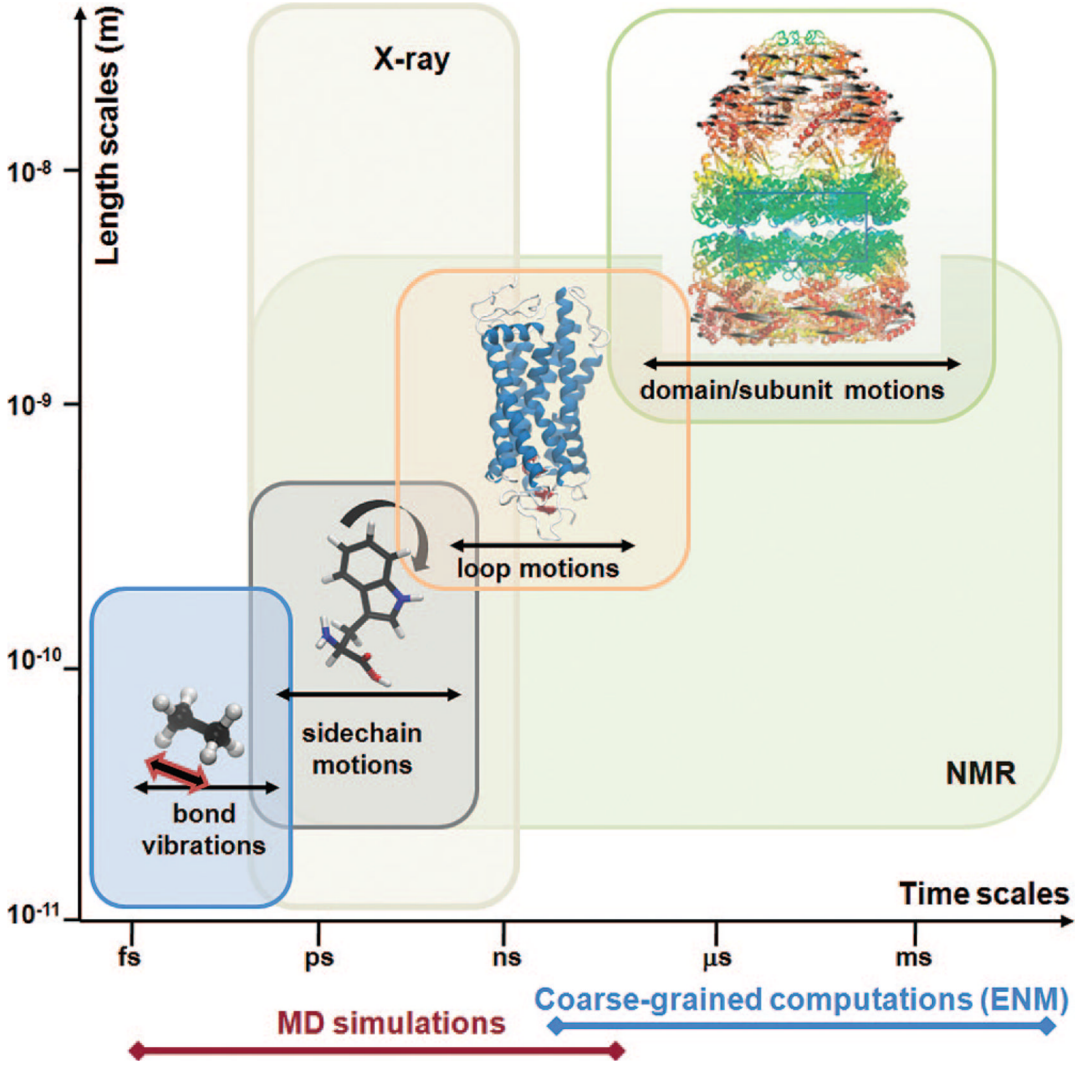

Figure 1.

Equilibrium motions of proteins. Motions accessible near native state conditions range from femtoseconds (bond length vibrations) to milliseconds or slower (concerted movements of multiple subunits; passages between equilibrium substates). X-ray crystallographic structures span length scales up to several hundreds of angstroms. Fluctuations in the subnanosecond regime are indicated by X-ray crystallographic *B*-factors. NMR spectroscopy is restricted to relatively smaller structures, but NMR relaxation experiments can probe a broad range of motions, from picoseconds to seconds, including the microseconds time range of interest for several allosteric changes in conformation. Also indicated along the abscissa are the time scales of processes that can be explored by MD simulations and coarse-grained computations. Molecular diagrams here and in the following figures have been generated using Jmol (http://www.jmol.org/), PyMol (http://www.pymol.org/), or VMD[35] visualization software.

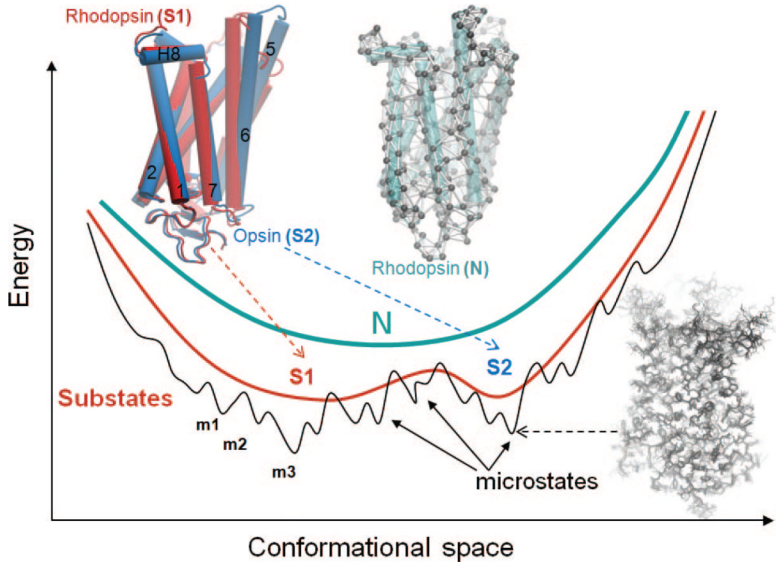

Figure 2.

Energy profile of the native state modeled at different resolutions. N denotes the native state, modeled at a coarse-grained scale as a single energy minimum. A more detailed examination of the structure and energetics reveals two or more substates (S1, S2, etc.), which in turn contain multiple microstates (m1, m2, etc.). Structural models corresponding to different hierarchical levels of resolution are shown: an elastic network model representation where the global energy minimum on a coarse-grained scale (N) is approximated by a harmonic potential along each mode direction; two substates S1 and S2 sampled by global motions near native state conditions; and an ensemble of conformers sampled by small fluctuations in the neighborhood of each substate. The diagrams have been constructed using the following rhodopsin structures deposited in the Protein Data Bank: 1U19 (N); 1U19 and 3CAP (S1 and S2); and 1F88, 1GZM, 1HZX, 1L9H, 1U19, 2G87, 2HPY, 2I35, 2I36, 2I37, 2J4Y, 2PED, 3C9L, and 3C9M (microstates).

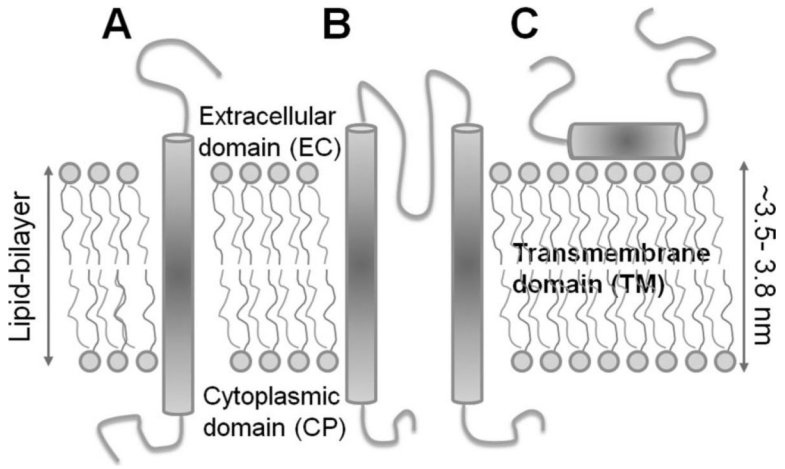

Figure 3.
Schematic representation of different types of integral membrane proteins. (A) Single helical TM protein (a bitopic membrane protein), (B) a polytopic TM protein composed of multiple TM elements (here two helices), and (C) an integral monotopic membrane protein.

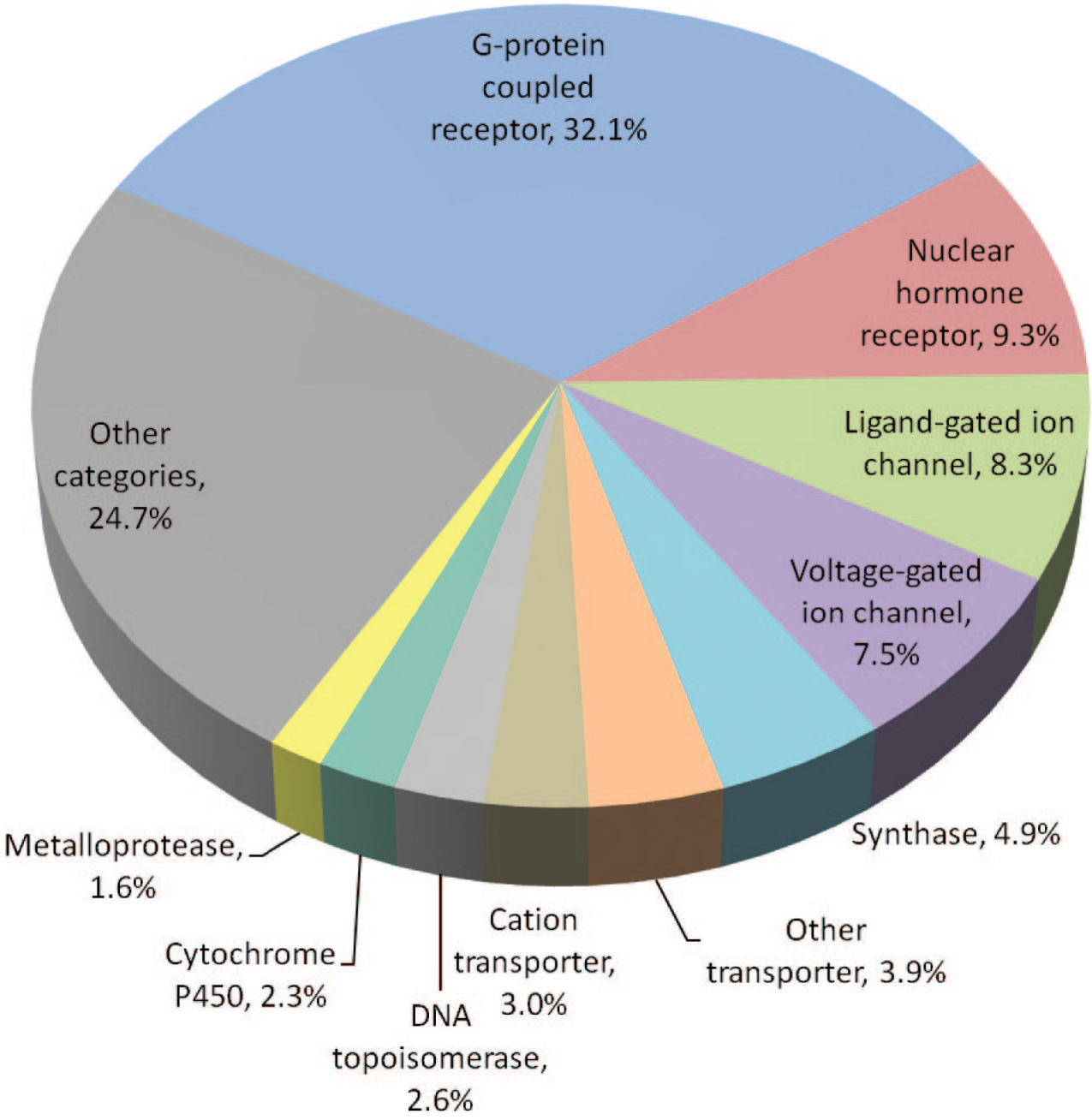

Figure 4.

Distribution of small molecule drugs based on the targeted molecular function. The distribution is shown for the top-ranking ten functional categories targeted by 965 FDA-approved small molecule drugs, excluding biotechnology drugs, nutraceuticals such as vitamins and amino acids, and those with uncertain targets. The top ten categories shown in the pie chart are associated with more than 75% of the drugs in the data set. The distribution is based on 1008 drug–protein associations. A given category is counted once if a given drug targets multiple proteins in that category.

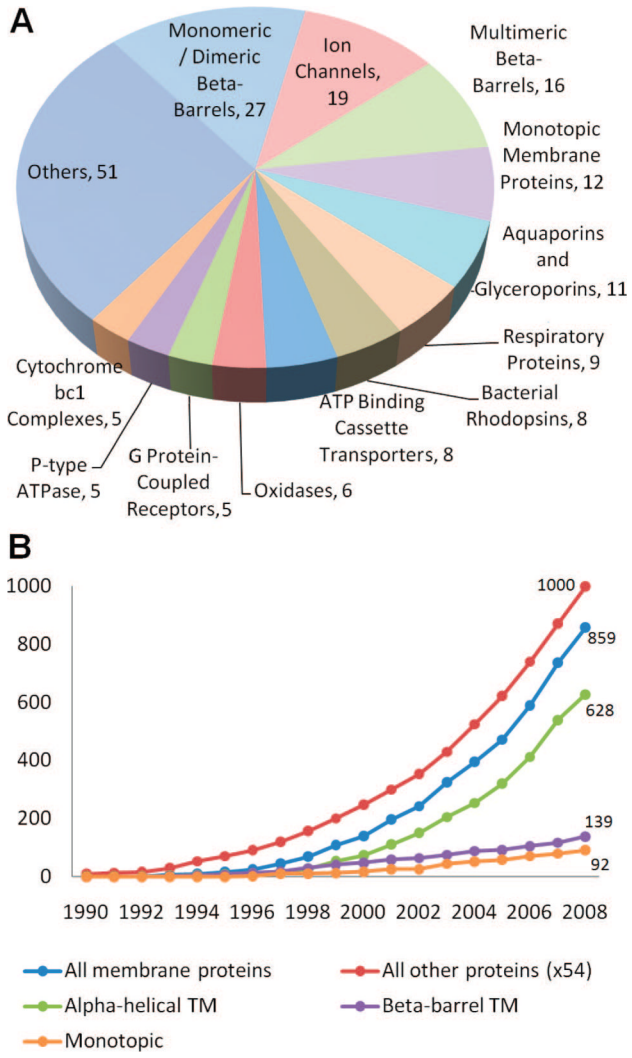

Figure 5.

PDB statistics for membrane proteins. (A) Structures of 182 unique membrane proteins are available in the PDB, as of January 2009. The pie chart displays the distribution of these structures among different functional/structural categories. (B) Growth of released membrane protein structures and other protein structures starting from 1990. Note that the number of "other" proteins is reduced by a factor of 54 in the curve, for display purposes. We also show the breakdown of membrane proteins into different structural groups: α-helical TM, *β*-barrel TM, and monotopic. An exponential growth with an R2 value of 0.99 is observed in the last ten years for both membrane proteins and all other proteins. The corresponding growth rates are 0.23 and 0.18, respectively; that is, they are higher for membrane proteins due to initiatives in that direction.

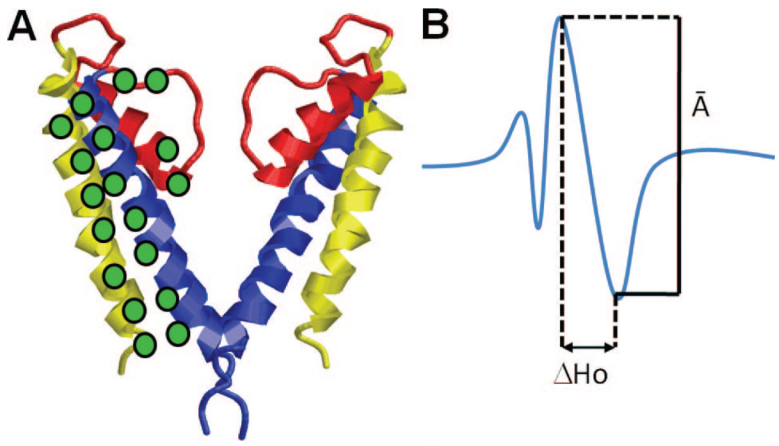

Figure 6.

Site-directed spin-labeling coupled with EPR illustrated for a potassium channel. (A) Molecular model of KcsA (omitting two of the four subunits for clarity). The green discs indicate the positions of the spin-labeled residues (probes) on TM1 (yellow), TM2 (blue), and the selectivity filter (red). (B) Measurement of the structural parameter from the spectral line shape of an EPR-spectrum. The amplitude (A^-) of the normalized central resonance line M = 0 and the mobility parameter ΔH_o (the peak-to-peak width at M = 0) are shown. Changes in two structural parameters are usually examined: (i) probe mobility (ΔH_o) and (ii) spin–spin interaction parameter W. Changes in the probe mobility, $\Delta\Delta H_o$, indicate rearrangements in tertiary or quaternary contacts, while the W parameter obtained from the ratio of the normalized amplitude spectra (A^-) in different forms reports changes in the intersubunit probe-to-probe proximity. Such measurements performed by Perozo and co-workers for the open and closed conformations of KcsA as a function of pH[157] revealed the coupled rigid-body rotations of TM helices TM1 and TM2 of the four subunits and the opening of the permeation pore (gating) induced by the concerted rotations of the TM2 helices while the selectivity filter remained practically immobile.

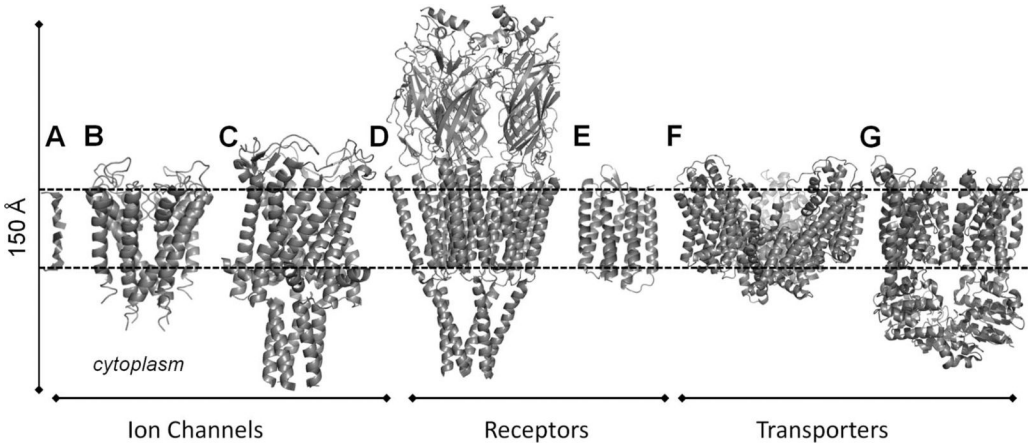

Figure 7.

Transmembrane proteins studied by NMA, considered in the present review: (A) gramicidin A, (B) KcsA, (C) MscL, (D) nAChR, (E) rhodopsin, (F) glutamate transporter (Glt_{ph}), and (G) BtuCD. The bilayer is indicated by the dashed lines. The ribbon diagrams were constructed using the respective structures 1NRU, 1K4C, 2OAR, 2BG9, 1L9H, 1XFH, and 1L7V available in the PDB.

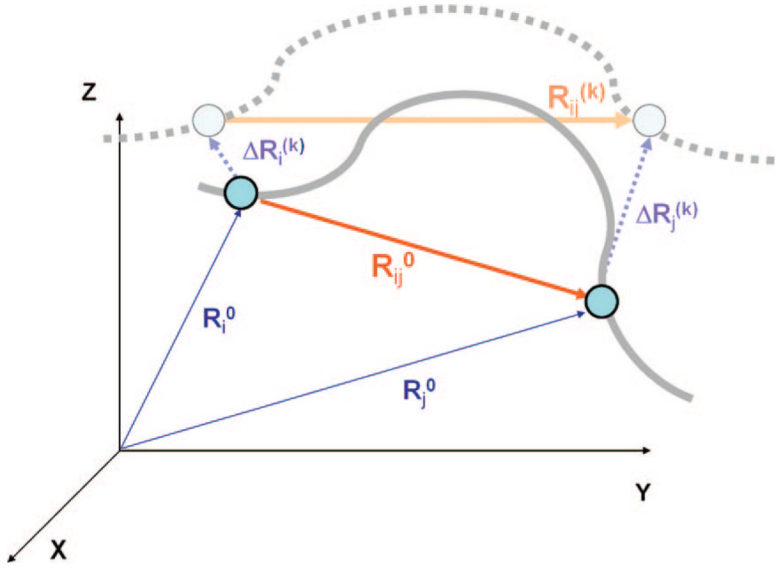

Figure 8.

Schematic view of interaction sites and their displacements. In the initial conformation, CG sites *i* and *j* are located respectively at $\mathbf{R}_i^0$ and $\mathbf{R}_j^0$, and the vector $\mathbf{R}_{ij}^0 = \mathbf{R}_j^0 - \mathbf{R}_i^0$ defines the distance vector between these sites. Upon displacement along mode *k*, the CG sites move to $\mathbf{R}_i^0 + \Delta\mathbf{R}_i^{(k)}$ and $\mathbf{R}_j^0 + \Delta\mathbf{R}_j^{(k)}$, and the distance vector becomes $\mathbf{R}_{ij}^{(k)}$. The solid gray line represents the structural details of the initial-state protein that are above the resolution of the coarse graining, and the broken gray line indicates the structure after a displacement along mode *k*.

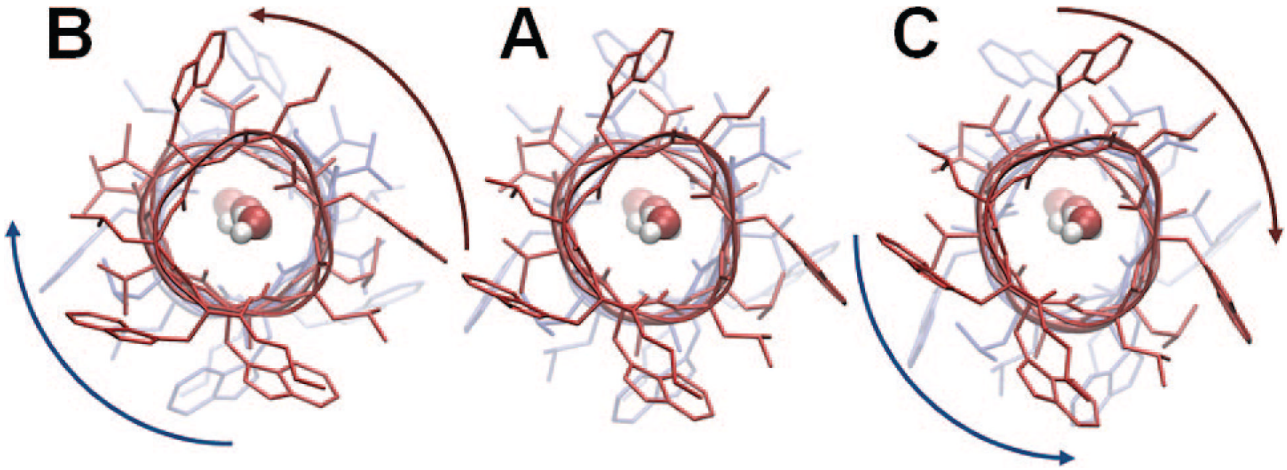

Figure 9.

Counter-rotations of the two helical dimers of gramicidin A, viewed from the EC side along the channel axis. This is the lowest frequency ANM mode of motion of the dimer. It is accompanied by a lateral expansion at the helical termini near the CP and EC regions. This mode was found to be crucially important for the initiation of the dissociation of the monomers needed for ion channel gating. Calculations were performed on an ANM server[78] (http://ignmtest.ccbb.pitt.edu/cgi-bin/anm/anm1.cgi), using the PDB structure 1JNO. Panel A displays the PDB structure, and panels B and C show two conformations fluctuating in opposite directions along the lowest frequency mode. Water molecules were placed inside the pore using Sybyl 8.3. (figure inspired by ref 192).

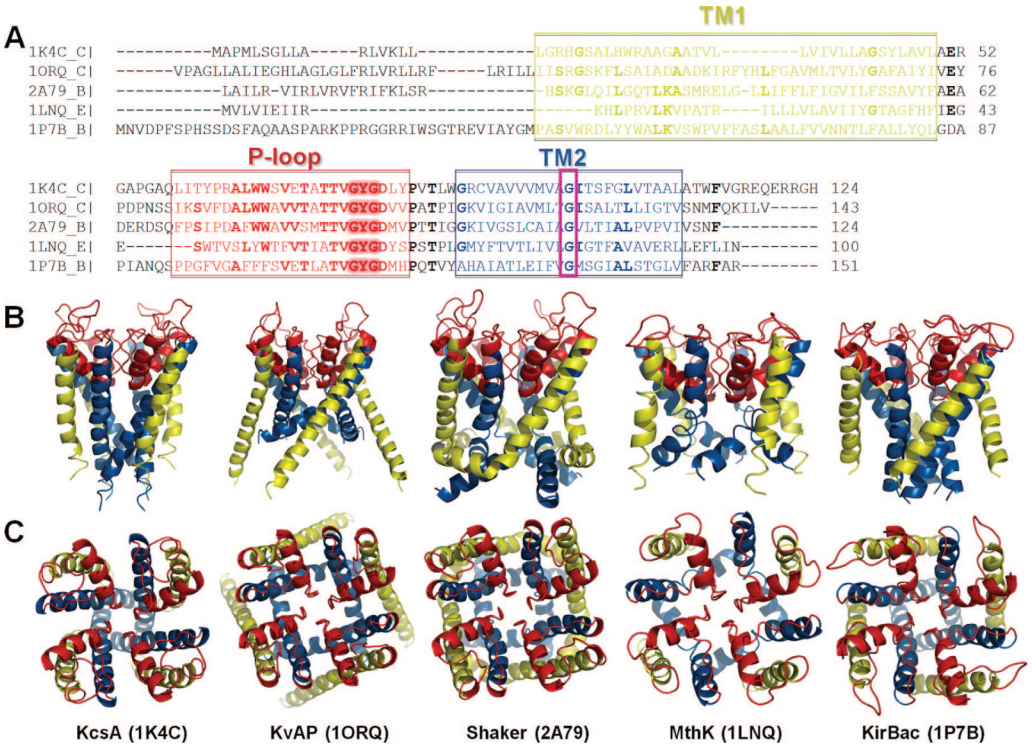

Figure 10.

Sequence and structure of the pore region of five structurally known K^+ channels. (A) Alignment of the pore region sequences. The regions corresponding to the helices TM1 and TM2 and the P-loop are indicated by the boxed green, blue, and red letters, respectively. The alignment was performed using ClustalW.[248] Fully or highly conserved amino acids are written in boldface. Two regions of interest are the signature motif GYG (highlighted) at the selectivity filter and the conserved glycine on TM2 (e.g., G83 in MthK) enclosed in a magenta box. (B and C) Structural comparison of the pore forming regions aligned in panel A. These are all tetrameric structures. The monomers contain either two TM helices (KcsA, MthK, and KirBac, with PDB ID's 1K4C, 1LNQ, and 1P7B, respectively) colored yellow (TM1) and blue (TM2) or six TM helices (KvAP and Shaker with PDB ID's, 1ORQ and 2A79, respectively). The helices S5 and S6 of KvAP and Shaker are equivalent to the respective helices TM1 and TM2 of the other K^+ channels and are displayed here, along with the P-loop region, colored red. The channels are viewed from side (B) and from the top (EC region) (B) (see ref 218 for more details).

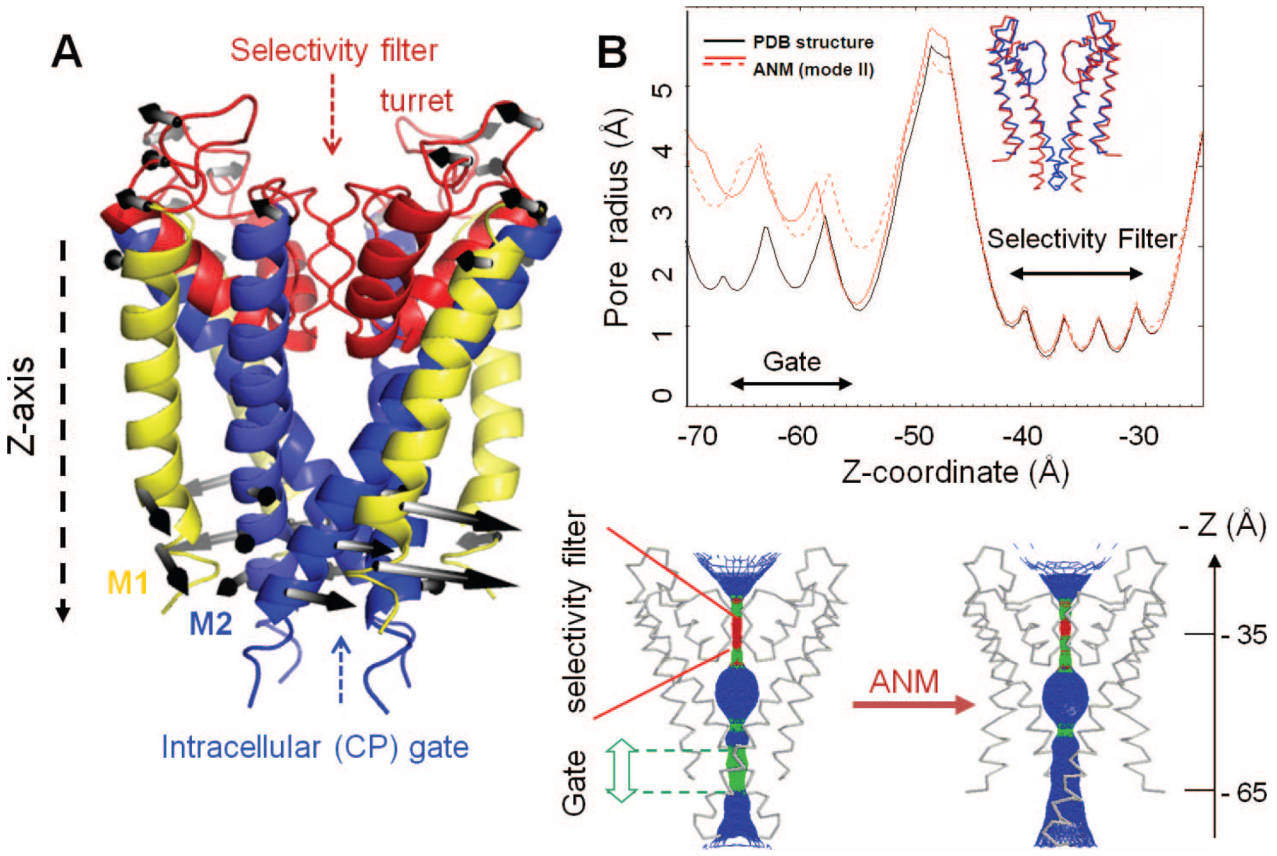

Figure 11.

Opening up of the potassium channel pore by the global twisting mode. (A) Ribbon diagram of KcsA illustrating the basic structural regions and the motion along the second slowest ANM mode. This is a global twisting (nondegenerate) mode that induces counter-rotations at the EC and CP regions, indicated by the white/gray arrows. (B) Top panel: The pore-radius profile as a function of the pore axis (*Z*-axis), calculated for the X-ray structure (black curve) and for two conformations visited by fluctuations in opposite directions along the global twisting mode (red curves). The inset shows the backbone trace of two of the monomers in the X-ray structure (blue) and the ANM-predicted conformation (red). The separation between the inner (TM2) helices at the gate is enlarged by about 1.5 Å. Bottom panel: A mesh-wire representation of the channel pore before (left) and after (right) reconfiguration along the second ANM mode. Color code: blue, radius > 3 Å; green, 3 Å > radius > 2 Å; red, radius < 2 Å. For visual clarity, only two monomers of the tetramer are shown (see ref 218 for more details).

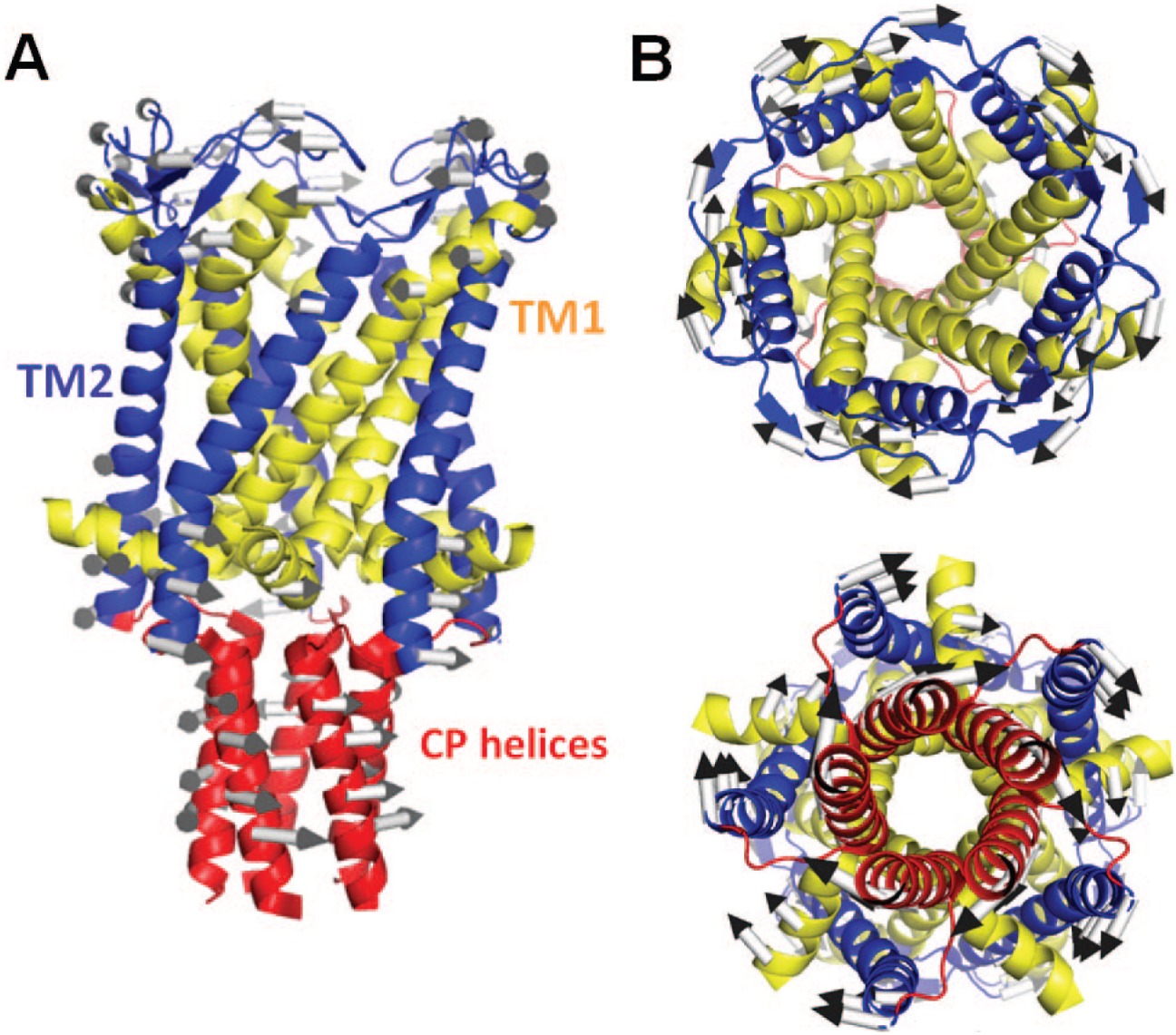

Figure 12.

Global dynamics of *M. tuberculosis* MscL predicted by the ANM. (A) Side view of the pentameric structure, and (B) views from the EC (top) and CP (bottom) regions. The TM helices are colored yellow (TM1) and blue (TM2), which are the inner and outer helices, respectively. CP helices are colored red. The TM and CP helices rotate in opposite directions in the slowest ANM mode. The directions of the arrows in panel B refer to the rotations as viewed from the EC and CP regions, hence their "apparent" rotation in the same direction. We also note that the structure fluctuates between two conformers where the TM helices and CP helices undergo counter-rotations, in either direction; that is, the arrows displayed in the figure represent one of the two opposite direction movements along this mode axis. The ribbon diagrams are generated using the structure (PDB ID: 2OAR) resolved by Chang et al.[117]

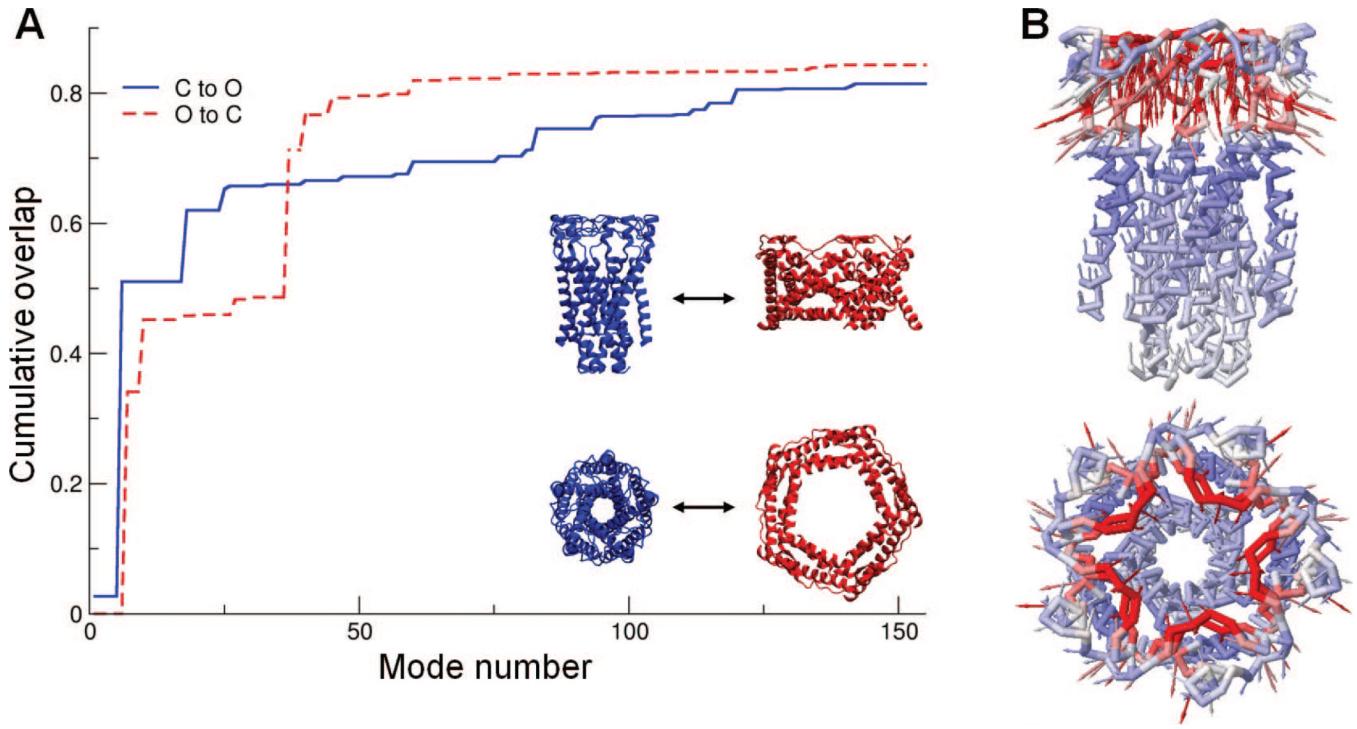

Figure 13.

Cumulative contribution of ANM modes to the structural change between the open and closed forms of MscL. The ordinate displays the cumulative overlap between the ANM modes (eigenvectors) predicted for the starting conformation and the targeted direction of structural change. ANM calculations were performed using either the closed (C) form (blue, solid curve) or open (O) form (red, dashed) as the starting substate. In either case, a cumulative overlap of about 0.8 is achieved by the top-ranking ~120 modes (less than 1/10th of accessible modes). Concrete (stepwise) contributions are made by the nondegenerate modes. The 2nd lowest nondegenerate mode accessible to the closed form (mode 6) is illustrated in panel B. This mode induces a contraction/expansion along the pentameric axis, mainly the portion close to the EC region, as seen from the side (top) and EC (bottom) views of the channel.

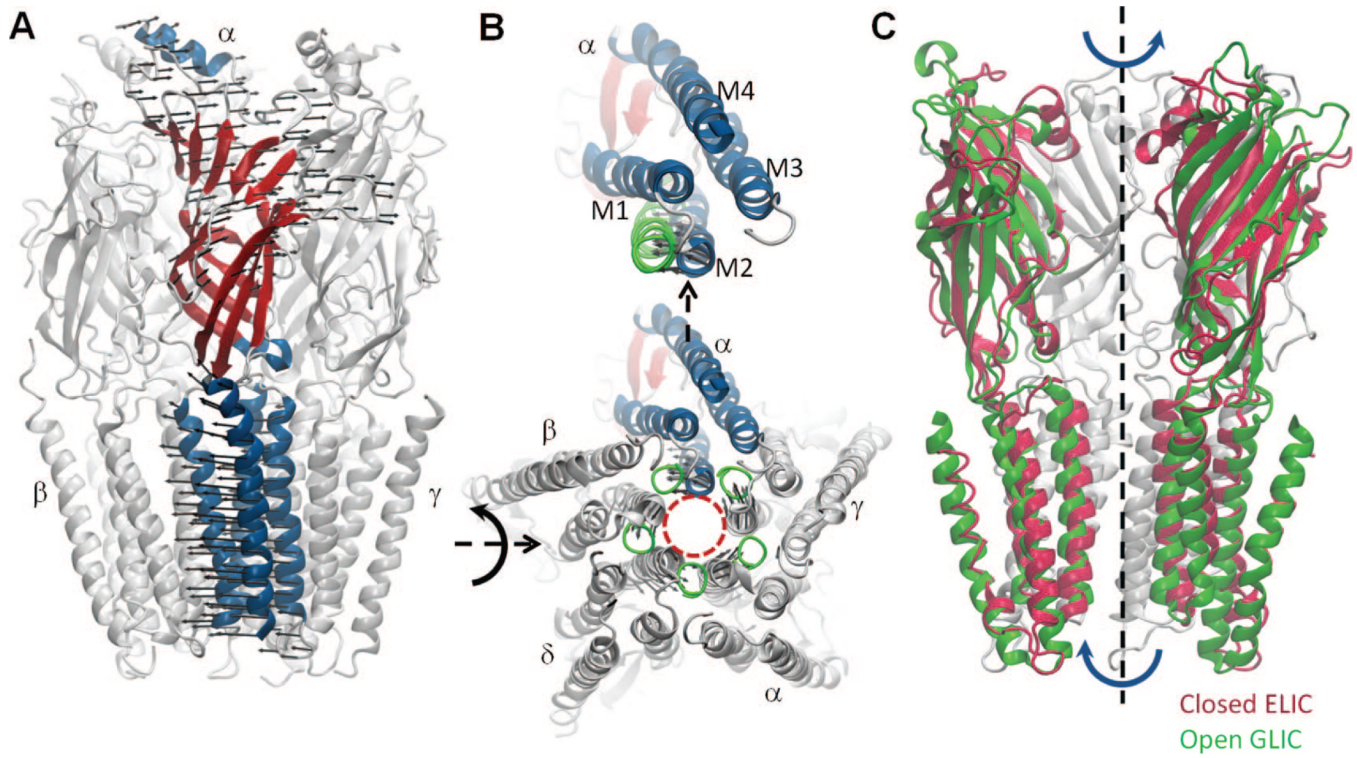

Figure 14.

Ligand-gated ion channel nAChR structure and dynamics. (A) Structure of the EC and TM domains of nAChR[265] (PDB ID: 2BG9). The secondary structure of one of the monomers (α) is colored to display the *β*-sandwich fold (red) of the EC domain and the four TM helices (M1–M4; blue) of the TM domain; and the remaining four monomers are shown in gray. The lowest frequency ANM mode induces a quaternary symmetric twist, as indicated by the arrows shown for monomer α. (B) CP end of TM domain (bottom) and close up view of one of the monomers (monomer α, colored) (top). Red dashed circle indicates the channel pore. Arrows indicate the collective movements of M2 helices along ANM mode 1. Green circles represent the CP end of the M2 helices after deformation along ANM mode 1. (C) Comparison of bacterial homopentameric LGICs ELIC (2VL0) and GLIC (3EAM) shows the contribution of this quaternary twist mode to the conformational changes involved in activation. One subunit (closest to the viewer) is omitted to display the channel pore.

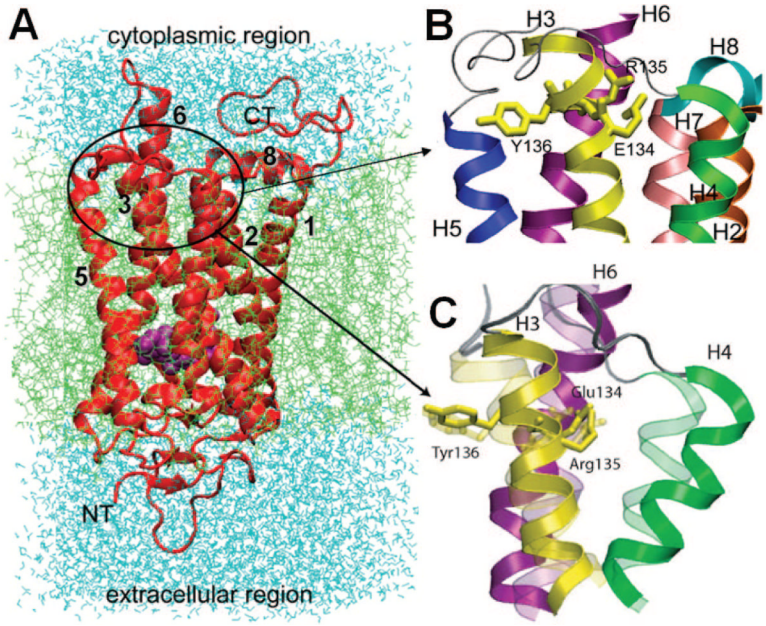

Figure 15.

Rhodopsin structure and its ERY motif at the CP region. (A) Ribbon diagram of the first rhodopsin structure determined by Palczewski and co-workers,[289] shown in a lipid bilayer. This is a seven TM helix structure, enclosing a chromophore (*cis*-retinal, shown in space filling, magenta). The C- and N-termini are labeled as CT and NT, along with some of the TM helices that can be distinguished clearly. Note that there is an eighth helix, at the CP region, that runs parallel to the membrane surface. (B) Enlarged view of the CP region containing the ERY motif (E134–R135–Y137) on the TM helix 3 (or H3), involved in G-protein recognition. (C) Reconfiguration of the ERY-motif-containing domain upon *cis–trans* isomerization of the retinal induced by light activation, suggested by an ANM analysis[85] of rhodopsin dynamics.

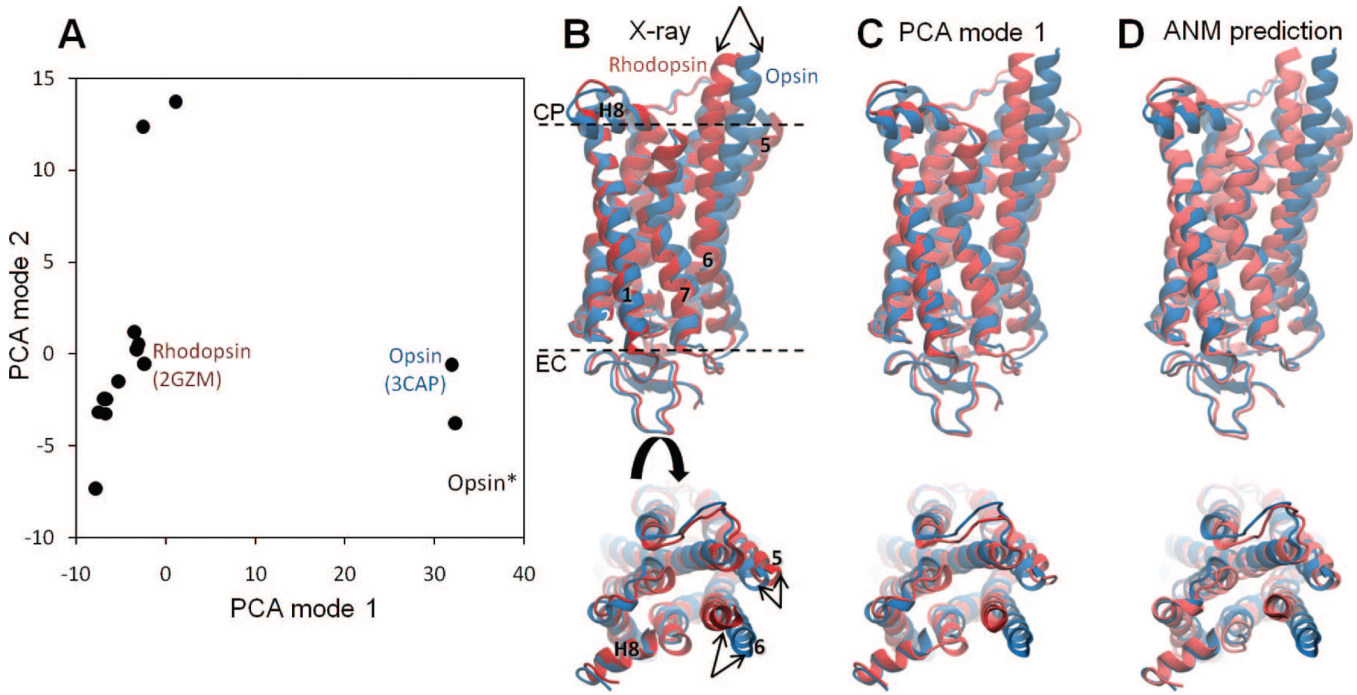

Figure 16.

PCA and ANM calculations for rhodopsin. (A) Distribution of 16 X-ray structures in the subspace spanned by the PCA mode directions 1 and 2. These respective modes account for 62% and 12% of the structural variability in the data set. The principal axis 1 differentiates the inactive and (putative) activated structures which are clustered in two distinctive groups, and the PCA axis 2 further differentiates between the structures in the cluster of inactive rhodopsins (B). Superimposition of experimentally determined rhodopsin and opsin structures, indicated by the labels on panel A. (C) Rhodopsin structure generated by deforming the opsin structure along the first principal mode, p_1. (D) Rhodopsin conformation predicted by deforming the opsin structure along the 20 lowest frequency ANM modes. The 14 rhodopsin structures in the analyzed set include, in addition to the ground state[289,307–311] and various photoactivated states, lumirhodopsin,312 bathorhodopsin,[313] 9-*cis*-rhodopsin,[314] photoacivated deprotonated intermediate,[309] and thermostabilized mutants.[310,315] These microstates are dispersed along the second principal axis. These calculations have been performed for the C_α atoms only; the remaining backbone atoms were reconstructed with the BioPolymer module of Sybyl 8.3 (Tripos). ANM calculations were performed using the relatively short cutoff distance of R_c = 8 Å, so as to release interhelical constraints.

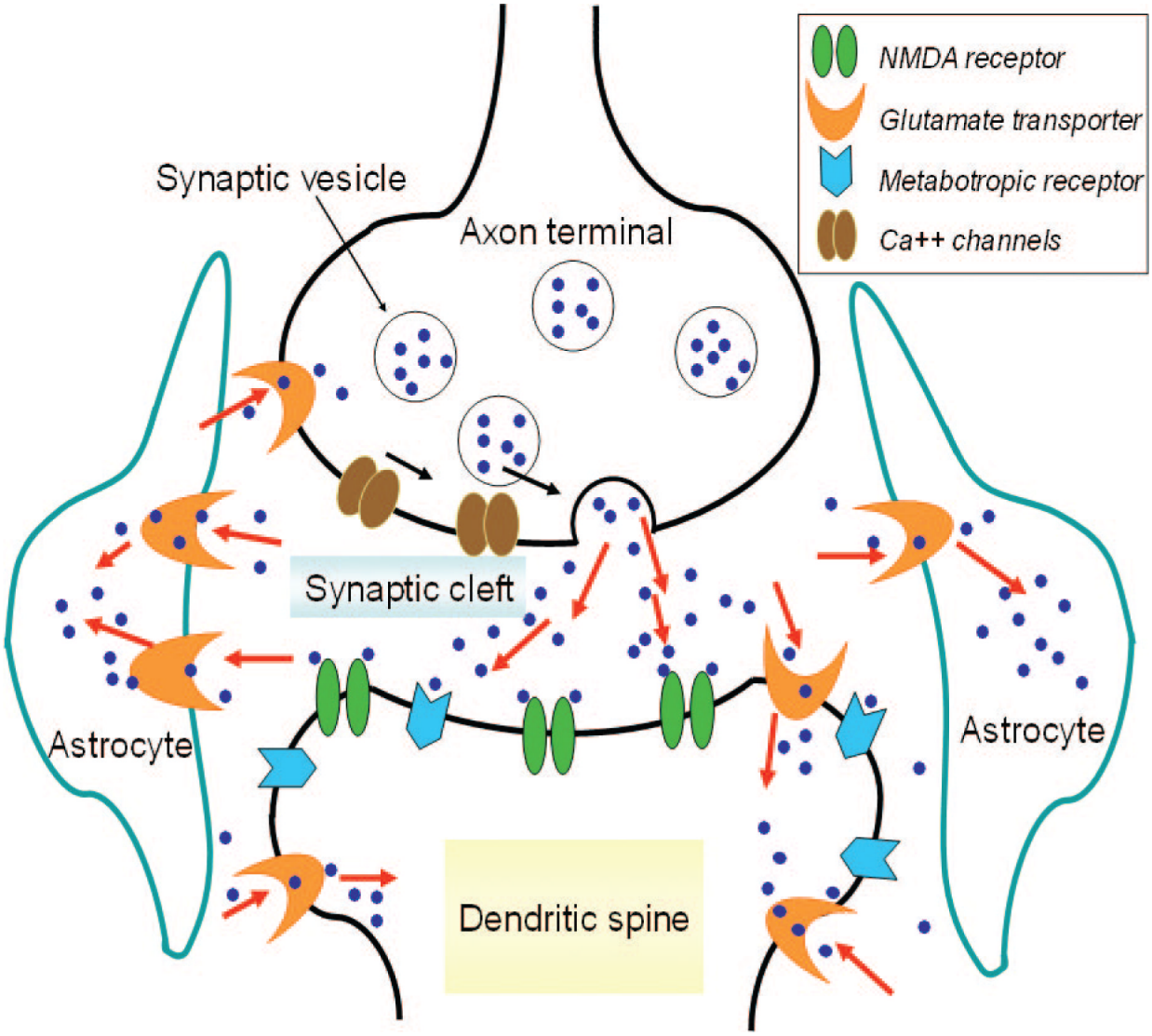

Figure 17.

Release, uptake, and reuptake of glutamate at an excitatory synapse. Upon arrival of an action potential at the presynaptic axon terminus, voltage-sensitive Ca^{2+} channels trigger the fusion of vesicles with the cell membrane to release glutamate molecules in the synaptic cleft. Glutamates bind and activate receptors on the postsynaptic cell membrane. Excess glutamate is cleared by glutamate transporters, which are more abundant and efficacious in the glia in the vicinity of the synapse.

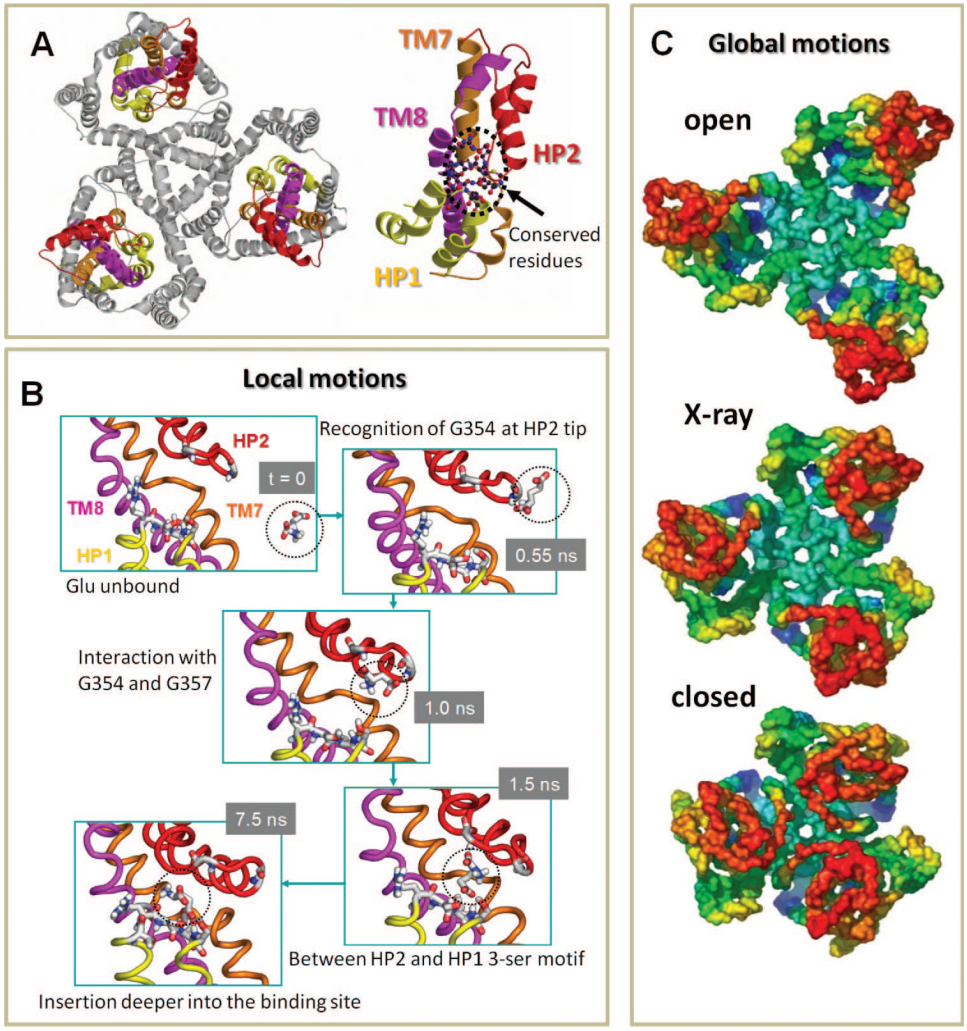

Figure 18.

Structure and dynamics of glutamate transporter. (A) The homotrimer, viewed from the EC region. The N-terminal region (TM1–TM6) is displayed in gray; the C-terminal core HP1–TM7–HP2–TM8 is colored yellow–orange–red–violet and labeled in the figure on the right. (B) Snapshots from MD simulations, illustrating the time-resolved recognition and binding events, starting from $t = 0$, where the substrate is in the aqueous cavity, up to $t = 7.5$ ns, where the substrate is sequestered at the binding site and remains therein for the remaining duration of the simulation, of ~20 ns.[326] (C) Symmetric opening/closing mode of Glt_{Ph}, as observed in ANM. The *middle* diagram displays the Glt_{Ph} structure viewed from the EC side (PDB: 1xfh); the *top* and *bottom* diagrams display the ANM-predicted open and closed conformations, respectively. In the X-ray structure, the basin is exposed to the EC aqueous environment, while in the closed form contacts between neighboring subunits occur (see, for example, the L34 loops colored red).

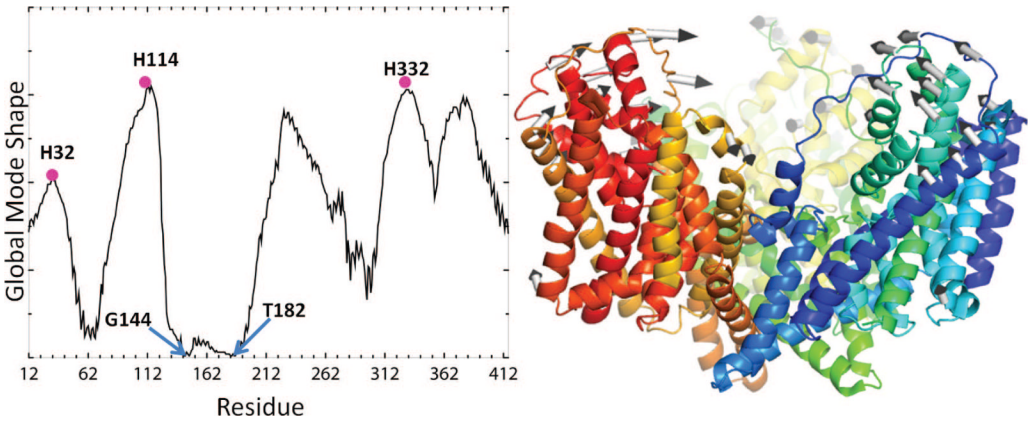

Figure 19.

Global dynamics of the aspartate transporter Glt_{Ph} predicted by the ANM. (A) Distribution of square displacements of residues, $(\Delta R_i)^2|_1 + (\Delta R_i)^2|_2$ (see eq 26), induced in the asymmetric stretching–gcontraction mode (a 2-fold degenerate mode). The same profile is induced in all three subunits upon superposition of these two lowest frequency modes, leading to a cylindrically symmetric reconfiguration. Peaks refer to the most mobile residues, and minima to the hinge centers (e.g., Gly144 and T182) controlling the concerted movements of the subunits. The large amplitude swinging movements of the extracellular histidines suggest a possible role in facilitating the attraction of the anions or engulfing them into the central basin. (B) Mechanism of motion in the first nondegenerate ANM mode (see also panel C in Figure 18). The arrows indicate the direction of the concerted movements of the three subunits (note that the third subunit in the back is lightly visible).

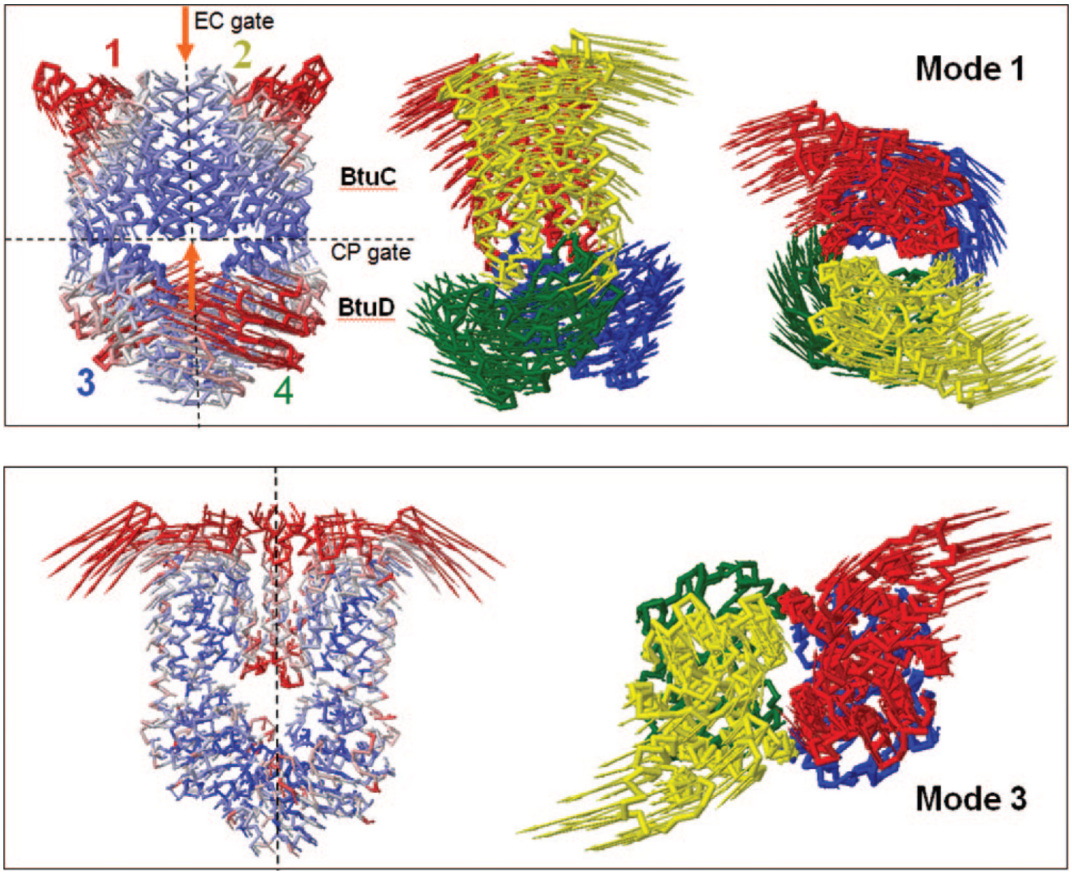

Figure 20.

Global dynamics of the ABC transporter BtuCD. Top and bottom panels display the collective motions of the tetramer in the ANM modes 1 and 3, respectively, recently examined by Weng et al. (2008). The color-coded diagrams on the *left* in both panels display the size of motions (red, most mobile; blue, almost rigid) induced in these modes. The other diagrams display the relative motions of the two TM domains (1, red; 2, yellow) of the BtuC dimer, and the two NBDs (3, blue; 4, green) of the BtuD dimer, that compose the BtuCD tetramer, viewed from the side (*middle* diagram in top panel) or from the EC region (*right* diagrams in both panel). The two gates (EC and CP gates) of the substrate (vitamin B_{12}) translocation pore are indicated by the orange arrows in the *left* diagram of the top panel.

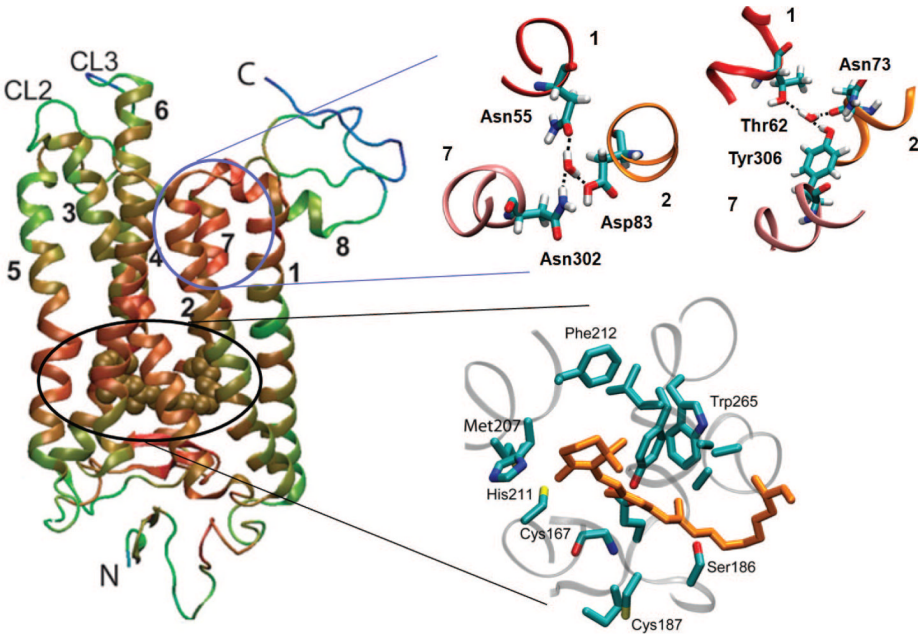

Figure 21.

Critical interactions near the chromophore binding pocket and CP ends of TM1, TM2, and TM3 in rhodopsin. The ribbon diagram on the left is color-coded (from red, least mobile, to blue, most mobile) by the RMSDs observed in the positions of residues during ANM-steered MD simulations of rhodopsin activation. Two regions enlarged on the right are distinguished by their highly constrained dynamics: the chromophore binding pocket and the CP end of helices 1, 2, and 7. The tight packing in the former region ensures efficient propagation of the local conformational strains (induced upon *cis* → *trans* isomerization of the retinal) to distant regions, including in particular the ERY-binding motif at the CP end of helices H3 and H6 (note the enhanced mobility at this region). Water molecules play an important role in stabilizing the CP ends of TM1, TM2, and TM7. For more details, see ref 85.

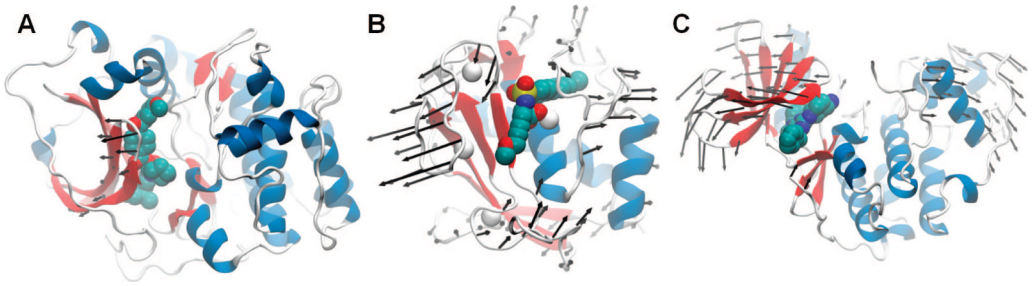

Figure 22.
Use of NMA in modeling protein–ligand interactions. Alternative conformations for the target protein were generated for (A) cAMP-dependent protein kinase382 (PDB ID, 1JLU; inhibitor PDB ID, 1REK), (B) matrix metalloproteinase-3383 (PDB ID, 1UEA; inhibitor PDB ID, 1HY7), and (C) cyclin-dependent kinase 2384 (complex PDB ID, 1G5S), by reconfiguring these target proteins along the global modes of motions indicated by the arrows. See text for details.

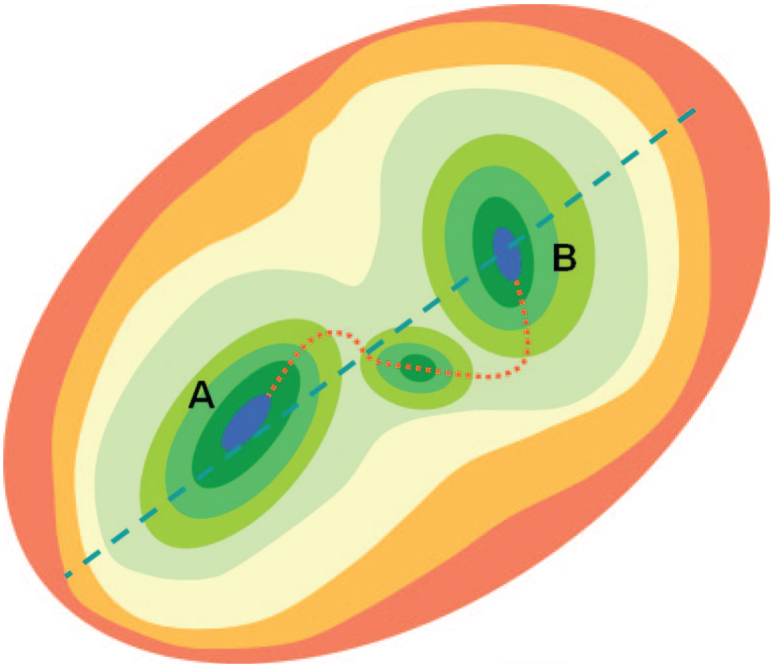

Figure 23.

Schematic representation of the energy landscape for two substates. The cartoon shows the putative free energy landscape around a conformational transition for a two-state system. Both conformations, A and B, are contained within a global free energy well, represented here as the outermost oval. The slowest mode of the well, indicated by the broken blue line, is expected to overlap with the transition between states A and B. Each stable conformation lies at the bottom of its own local well. The transition between states (red dotted line) is expected to proceed along the slowest local mode in the vicinity of each end point. The slow modes accessible to the metastable intermediate conformation between the end points provide further information on the pathway near the transition point.