

Perspective

Are Protein Shape-Encoded Lowest-Frequency Motions a Key Phenotype Selected by Evolution?

Laura Orellana

Protein Dynamics and Mutation Lab, Department of Oncology-Pathology, Karolinska Institute, Solnavägen 9, 171 65 Solna, Sweden; laura.orellana@ki.se

Abstract: At the very deepest molecular level, the mechanisms of life depend on the operation of proteins, the so-called “workhorses” of the cell. Proteins are nanoscale machines that transform energy into useful cellular work, such as ion or nutrient transport, information processing, or energy transformation. Behind every biological task, there is a nanometer-sized molecule whose shape and intrinsic motions, binding, and sensing properties have been evolutionarily polished for billions of years. With the emergence of structural biology, the most crucial property of biomolecules was thought to be their 3D shape, but how this relates to function was unclear. During the past years, Elastic Network Models have revealed that protein shape, motion and function are deeply intertwined, so that each structure displays robustly shape-encoded functional movements that can be extraordinarily conserved across the tree of life. Here, we briefly review the growing literature exploring the interplay between sequence evolution, protein shape, intrinsic motions and function, and highlight examples from our research in which fundamental movements are conserved from bacteria to mammals or selected by cancer cells to modulate function.

Keywords: protein dynamics; evolution; intrinsic motions; elastic network models



Citation: Orellana, L. Are Protein Shape-Encoded Lowest-Frequency Motions a Key Phenotype Selected by Evolution? *Appl. Sci.* **2023**, *13*, 6756. <https://doi.org/10.3390/app13116756>

Academic Editors: Robert Jernigan and Domenico Scaramozzino

Received: 7 February 2023

Revised: 17 May 2023

Accepted: 19 May 2023

Published: 1 June 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. From the Structure–Function Paradigm to Structure–Motion–Function

Over 60 years ago, Anfinsen’s postulate that “the native secondary and tertiary structures are contained in the amino acid sequence itself” [1] laid out the foundations of the central dogma of structural biology, i.e., that the sequence of a protein contains the information required to adopt a defined 3D-structure and, hence, function (see historical overview in [2]). This so-called structure–function paradigm was formulated during the time when biomolecular crystallography was flourishing. According to Martin Karplus, X-ray crystallography created “the misconception . . . that the atoms in a protein are fixed in position” [3]. This view is also shared by cryo-EM pioneer Joachim Frank, who wrote that “the idea of “a” molecular structure has been largely created by X-ray crystallographic practice” [4]. As a consequence, a static view of proteins, in which one sequence folds into a unique “native conformation” responsible for function, became prevalent. Nevertheless, an alternative, dynamic view of proteins as an ensemble of conformations, more akin to the principles of physics, had been proposed long before by Pauling, Landsteiner, and others in the 1930s [5]. Fast forward in time to our days, and this early dynamic vision appears prescient. As our technology to capture proteins in action evolved (NMR, cryo-EM, etc.), it became clearer every day that proteins do not fold into a single static “native” structure, but are rather dynamic machines in continuous motion that explore complex and rugged energy landscapes [6], transitioning between multiple meta-stable minima. Such transitions encompass a wide hierarchy of time and length scales—from picosecond atomic fluctuations to microsecond or millisecond allosteric changes or breathing motions—and, importantly, are instrumental for proteins to sense and respond to environmental signals like ions or ligands [6–8].

Protein motions not only mediate or execute biological work—channel gating, ion pumping, transport, etc.—but also reshape interactions with other partners. Therefore, they

are central for molecular recognition [9–11], no matter whether it involves conformational selection or induced fit [12,13]. Even eminently local processes such as enzyme catalysis can involve dynamic changes such as side chain fluctuations or the unfolding of binding sites [14–16]. For intrinsically disordered proteins, flexibility is so extreme that the classical concept of a discrete number of well-defined native 3D shapes or conformers becomes almost meaningless; they can only be statistically described as ensembles of interconverting conformations [17,18]. Nevertheless, a majority of proteins fall in the middle ground between perfect rigidity and chaotic disorder, a boundary where discrete rigid domains or subunits exquisitely rearrange in response to signals. Cooperative motions, allosteric propagation, and large-scale conformational changes spontaneously emerge from this frontier of harnessed flexibility to create function, as pioneering work by Dorothee Kern showed [16].

Back in 1987, Elber and Karplus first noted the similarity of MD fluctuations with evolutionary changes across the globin family [19], inaugurating a fruitful line of evolutionary and structural dynamics comparisons to this day. Since then, structural data have grown exponentially, and Elastic Network Models (ENMs) [20–23] have revealed that such fluctuations are largely defined by molecular shape and determine functional motions. Overall, this has led to a new structure–motion–function dogma, where molecular shape determines intrinsic motions, and motions make function, a concept increasingly supported via cryo-EM ensembles [24,25]. Therefore, it is time to ask: if molecular motions mediate function, are they maybe a key object of evolutionary selection? Here, we briefly review evidence from structural biology and ENMs research, that points to shape-encoded motions as an essential matter for evolution.

2. ENMs Overview and the Surprising Accuracy of Shape-Encoded Harmonic Motions

A central problem in the study of protein dynamics has always been the difficulty of capturing motion, i.e., fully sampling conformational spaces. Protein flexibility is challenging to trap, describe, and predict, both experimentally and computationally [26]. Despite advances in hardware and algorithm parallelization, fully atomistic Molecular Dynamics (MD) simulations are still only feasible for ns– μ s timescales and middle-sized proteins. To gain insight into the mechanisms of bigger sub-mesosopic systems or the slow large-scale transitions associated with biological function, the physical description needs to change accordingly to lower-resolution Coarse-Grained (CG) models. Among the plethora of CG methods to model the dynamics of proteins, ENMs stand out as possibly the most simple and powerful, considering the balance between their minimal computational cost and striking predictive power. ENMs can be described as the CG flavor of Normal Mode Analysis (NMA), a classical mechanics technique used since the 1940s–1950s to analyze the vibrational spectra of simple molecules [27,28]. Soon after the first MD simulations, in 1982–1983 [29–33], NMA was applied for the first time to proteins to gain insight into their near-equilibrium dynamics. Instead of numerically solving Newton's equations as MD does, NMA assumes the harmonicity of the system around an energy minimum and, thus, through diagonalization of the mass-weighted Hessian matrix, allows the computation of a unique analytical solution, i.e., a set of linearly independent Normal Modes (NMs) (see details in [21,34]). NMs are a series of eigenvectors (v_i) ordered by their eigenvalues or frequencies (λ_i), that describe the natural motions of the system. Importantly, the first 5–10 ones, the so-called lowest frequency, “soft” or “slow” modes, capture the largest amplitude, more collective, and energetically “easiest” movements, which usually coincide with the experimentally and biologically relevant ones, as we will discuss below.

Despite its simplicity versus MD, NMA was still computationally heavy for large systems, as it required energy minimization and significant memory resources for matrix diagonalization. Inspired by early “random networks” and “beads-and-springs” polymer models developed by Flory and Rouse [35,36], ENMs took the simplification of NMA one step further, replacing detailed physical force fields with a minimalist representation of proteins as networks of residue nodes connected with elastic springs, devoid of chemical

or sequence information. Moreover, the system was assumed to be already at a minimum, skipping energy minimization. The first ENM [37] was still an all-atom model but with a simple pairwise Hookean potential: the native structure was defined as the minimum, and detailed interactions were replaced with a squared potential and a uniform constant within a cutoff. Shortly after, Bahar's one-dimensional Gaussian Network Model (GNM) [38] introduced the coarse-graining of structures to the C α trace, and finally, the Anisotropic Network Model (ANM) [39] combined Tirion's 3D-model with GNM coarse-graining, becoming the basis for most ENM methods nowadays [22,40]. The similarity of the motions described using coarse-grained ENMs with the atomistic Tirion's model, and of Tirion's with classical NMA based on accurate molecular potentials, was initially puzzling. How can such minimal one- or two-parameter models reproduce the vibrational properties of a complex macromolecule? The answer lies in the fact that soft modes involve coherent motions of large groups of atoms, and thus are mostly defined by the overall mass/domain architecture. For that matter, CG and atomistic mappings are nearly equivalent.

ENM–NMA can have apparent simplicity—with “toy” ad hoc force fields and the naïve assumption that structures are in an energy minimum—but it is often unsurpassed in the prediction of experimentally observed large-scale conformational changes (Figure 1, center). There have been endless studies comparing ENMs with functional transitions between bound/unbound, active/inactive and open/closed pairs derived from X-ray conformers, NMR ensembles, etc., which show that the lowest-frequency modes are indeed both biologically and functionally relevant [41–44] and can unravel complex allosteric mechanisms [45], even for subtle transitions such as those seen in GPCRs [46–48]. Protein conformational changes often involve large rigid-body motions, e.g., domain swapping, hinge-bending, or shear movements, which are strikingly well described via a small number of ENM modes [49–51]. An early study on the first database of molecular motions, MolMov [52], determined that 95% of experimentally observed transitions can be described using just a couple of soft ENM modes. Further benchmark studies have confirmed that large-scale motions also coincide with the collective modes extracted from MD simulations or experimental ensembles [53–57] via Principal Components Analysis (PCA, see [58–60]). Systematic comparison with MD of representative meta-folds in the MODEL database as well as with experimental data [61,62] confirmed that ENMs are extremely robust to spring definitions and perform exceedingly well in predicting large-scale transitions, occasionally surpassing MD simulations.

Nevertheless, as often happens with CG models, a major weakness of ENMs is the lack of a consistent and universal consensus on force-field parameterization, i.e., the functions used to determine the “springs” connecting different residues or “beads”. This has both positive and negative aspects. On one hand, although ENMs can predict the preferred directions for conformational change, the time and length scales of the motions (i.e., the magnitudes of the eigenvalues) are usually arbitrary. On the other, and paradoxically, this weakness reflects their major strength: ENMs are determined by protein shape, topology, and local packing density, and are thus insensitive to fine details. Despite these shortcomings and their dramatic simplicity, soft ENM modes are surprisingly accurate at predicting anharmonic, far-from-equilibrium transitions [20,40]. Together with the lack of a solvent and thus damping, this was initially a major point of controversy, questioning the validity of both NMA and its CG approximation [63]. What is the time and length scale of NMs? How can harmonic NMs capture anharmonic, damped and slow transitions over high energy barriers? It has been argued that proteins oscillate around the equilibrium, with energy increasing as they stretch along NMs' directions. This could elegantly agree with a dynamical systems perspective, as the Kolmogorov Arnold Moser (KAM) theorem assures the persistence of quasi-periodic motions under small perturbations [21,64]. Under this view, NMs would define major directions around a potential well, that hold relatively far from equilibrium. Following these, the high energy states reached would be further stretched and stabilized by different ligands or signals capable of “tipping” the free energy landscape (the so-called pre-existing equilibrium model [65,66], experimentally observed

in enzymes [16]). Already in the 1990s, MD studies showed that indeed, the energy surface probed via simulations is well-approximated by a rescaled version of the harmonic potential [67,68]. Recent work has related anharmonicity to mode collectivity: low-frequency modes that are collective enough, remain harmonic even for large displacements and better correlate with experimental transitions [69]. The power of ENMs to explore the boundaries of free energy minima is thus being more and more recognized, to the point that they are now used to enhance sampling via MD [70]. Regarding the timescales question, it is clear that NMs cover all the protein motion timescales, from MHz (μ s) large-scale motions to 1–10 THz (ps) backbone/atomic vibrations. However, the actual NM eigenvalues are typically meaningless and need rescaling, with few exceptions like the nearest-neighbors ED-ENM model [54]. Apart from this arbitrary amplitude of single modes, ENM–NMA tends to spread variance at higher frequencies in comparison to MD Essential Dynamics (ED) modes [58,71], probably as a consequence of the absence of damping. Our ED-ENM model [54], developed from database-wide comparisons with MD force fields, attempted to solve these issues by fitting spring functions not only to predict conformational changes but also to obtain realistic amplitudes for the eigenvalues and their distribution (i.e., the actual time and length scales in solution). This study also revealed that even extremely simple ENMs, just connecting the first three neighbors in the peptide chain, can predict MD and experimental flexibility, which critically depend on peptide backbone topology and local cohesiveness.

In brief, despite their many weaknesses—inconsistent parameterization, arbitrary time and length scales, lack of damping—the ability of ENMs to track functional large-scale motions—regardless of CG levels, spring definitions, or any sequence or local details—is stunning. Precisely in this fact lies the greatest physical insight they reveal: that proteins' overall packing, local connectivity, and shape determine intrinsic collective motions that poise them for function. These motions hold far beyond equilibrium and also across extremely long evolutionary scales, as we will discuss now.

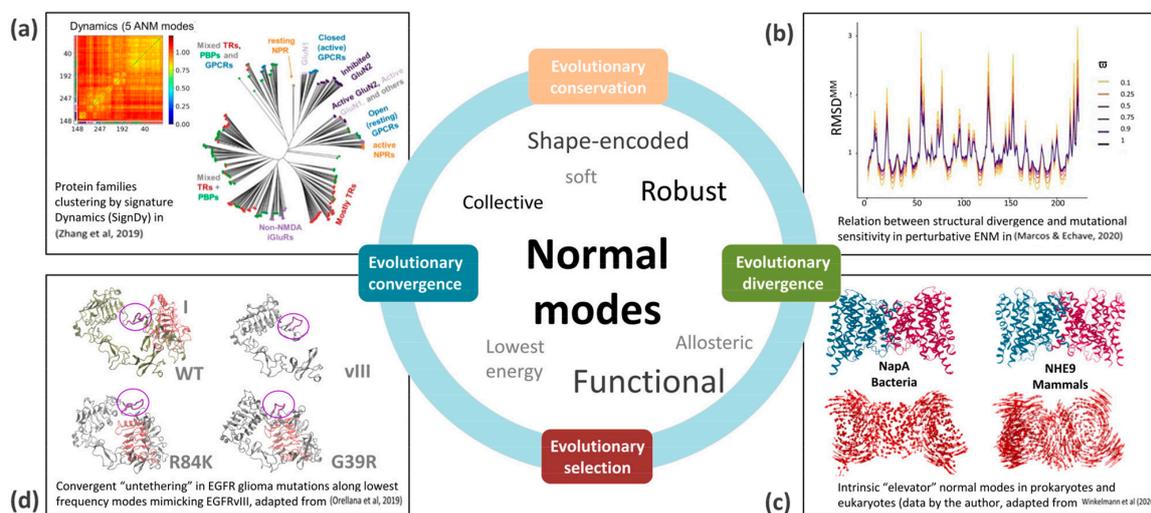


Figure 1. Shape-encoded ENM Normal Modes (NMs) and protein dynamics evolution examples from recent literature. **(a)** Signature Dynamics (SignDy) allows to build dynamics-based dendrograms comparable to those derived from sequence and structural similarity; see Ref. [72]. **(b)** Perturbative ENM suggests structural divergence relates more to mutational sensitivity (RMSD^{MM}) than selection (ω), which only deepens the profiles. See details in Ref. [73]. **(c)** Prokaryotic–eukaryotic conservation of NMs coupled to function and **(d)** Mutational convergence to favor an NM transition towards an oncogenic intermediate characterized by the exposure of a cryptic epitope (purple circle). See also the discussion in Section 4 and further details in Figure 2 and Refs. [74,75], respectively. Images (a,b) have a Creative Commons Attribution License and (c,d) are adapted by the author from her work.

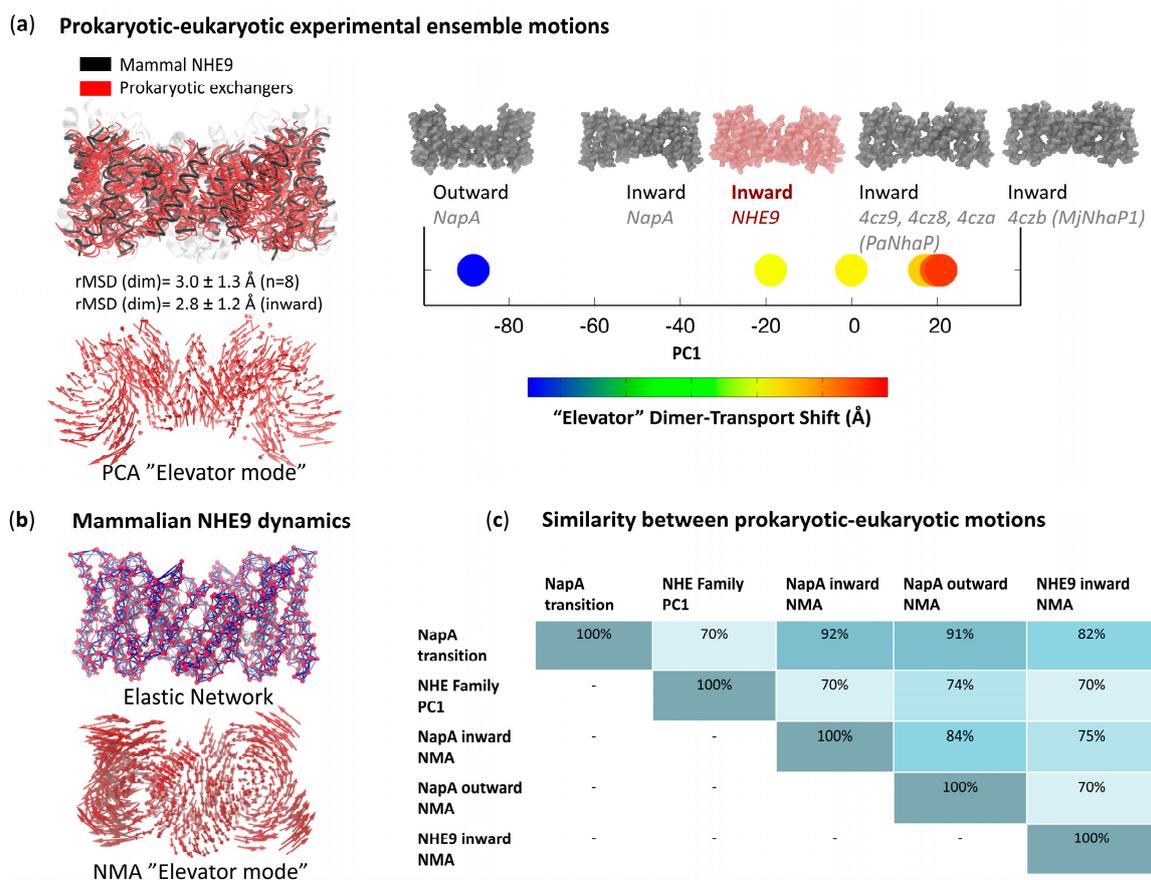


Figure 2. A closer look at CPA exchangers’ “elevator modes” conserved from bacteria to mammals. **(a) Left:** Core alignment between a mammalian exchanger, NHE9 (black) and distant bacterial homologs NapA, PanNhaP and MjNhaP1 (sequence identity $\approx 20\%$). The first principal component (PC1) of this ensemble of $n = 8$ structures renders the well-known “elevator-like” motion that distinguishes outward and inward states. **Right:** Projections onto PC1 of the experimental ensemble track the conformational inward-to-outward pathway and assigns the conformational status of the solved structures along it. **(b)** ENM of the mammalian NHE9 structure and derived “elevator-like” NM. **(c)** Similarity between NMs, PC1 and the prokaryotic NapA transition are all above 70%, despite the low sequence identities. Overlaps between vectorial spaces shaded in gradient; note that overlaps around 20% are considered random and from 40–50% significant. Adapted from figures and data by L. Orellana in Ref. [74], under the Creative Commons Attribution 4.0 License.

3. Lowest-Frequency Modes and Evolution

At the macroscopic level, we can easily appreciate how form, biological motion, and function evolve together under the laws of physics, shaping animal and plant morphologies [76]. Evolution seems to select the shapes best suited to perform functional motions. In the molecular world, if we assume the structure–motion–function paradigm, i.e., from motion comes function, it just follows to wonder whether evolution is selecting dynamics and resulting function rather than sequence or shape. Is there evidence of direct evolutionary pressure on protein motion? It is in this arena—where molecular evolution meets protein biophysics—that conformational dynamics becomes central [77]. Lowest-frequency modes allow for quantitative comparisons of the dynamics linked to function between similar cores [78], which are shedding new light on these questions.

Back in the 1980s, as soon as enough structures accumulated in the Protein Data Bank, it emerged that homologous proteins share similar folds, but this similarity wanes with increasing evolutionary distance [79,80]. Still, in practice, proteins with sequence similarities as low as 20% can display identical cores. The space of protein sequences is

known to be much larger than that of structures, close to optimal [81] and restrained by the length, stability, and topology of each fold [82]. Importantly, from this fact, it also follows that structural folds, i.e., protein shapes, are highly robust against mutations. What about conformational spaces? ENMs have revealed that each structure preferentially samples a limited set of elemental motions; the shape determines the conformations/motions, and the motions define the function. Being defined by global shape, soft modes are also incredibly robust to perturbations like mutations [83] or local structural features, and therefore hold across protein families and even remote homologs. Hence, when two sequences have low but sizeable sequence similarity, they often share a common core, motions, and probably function [84]. Moreover, proteins sharing one similar conformation often share other conformations, i.e., their conformational spaces are conserved, a concept exploited to predict new conformers or model conformational changes [85]. Therefore, we could argue that, in the same way the sequence space is bigger than the structure space, the structure space is bigger than the motion space—and this inversely relates to fold and function robustness.

Based on mounting evidence from ENMs [20] and parallel studies on residue flexibility [86], protein global dynamics has been suggested to be maximally conserved versus sequence and structure. Nevertheless, the degree of conservation of conformational spaces as well as the contributing factors are unclear. Due to the entanglement of function, motion, and shape, together with protein biophysical and evolutionary constraints, the issue is intensely debated [87–89]. There are two central questions to be addressed: Is it function that primarily drives the conservation of dynamics? Or is it due to physical constraints such as stability, topology, local packing, etc., or properties like mode energies or robustness? What about evolutionary constraints such as population sizes, mutational rates or bias? In other words: are soft modes conserved because they are functional or because they are energetically “easy” and robust? Probably, the truth is in the middle.

Evidence for direct evolutionary pressure on normal modes is still scarce, as quantitative comparisons of functional dynamics are relatively recent [78]. It has been proposed that there is negative selection against the divergence of functionally important modes, while other studies suggest that they are conserved just because they are more robust to mutational perturbations (Figure 1a,b). Soon after ENMs were developed, it became evident that proteins with similar architecture shared similar motions [90]. Early studies on the evolution of soft modes, led by Ortiz and colleagues, focused on how structural cores modify their shape across homologous proteins [91–93]. These pioneering works revealed significant similarity in the conformational ensembles explored within a superfamily and the soft modes, i.e., proteins seem to evolutionarily diverge along soft modes or, vice versa, protein topology constrains evolutionary divergence. In parallel, Echave also showed that the lowest-frequency modes are conserved in homologous proteins [94], and there is a significant correlation between mode collectivity and its conservation [95]. The conservation of lowest-frequency modes is apparent in residue fluctuation patterns, which can be easily aligned for homologous proteins [96]. Some studies have also pointed out that protein sites evolve at different rates depending on properties such as their solvent accessibility, packing density, and flexibility [97,98]. In general, there is an inverse relation between local flexibility and evolutionary rates [99] i.e., exposed and flexible loops are less conserved than cores or rigid regions [100], which can act as hinges for global motions. Consequently, ENM analyses show clear correlations between sequence evolution and structural dynamics, especially relevant for hinge regions [100,101]. These rigid regions are so critical that hinge migration has been proposed as a mechanism for protein evolution [102]. Moreover, cancer and disease-related mutations tend to focus on hinge-like areas [103,104]. Therefore, ENM dynamics is a key predictor of functional impact for point mutations [105,106] as well as for insertions and deletions [107], further discussed below.

Importantly, even in the case of random mutations, structural changes correlate with the lowest frequency modes [108], as happens also for ensembles of the same protein determined in different experimental conditions [109]. Perturbative ENMs indicate that the con-

servation of soft modes might arise precisely from their robustness against mutations [110] and, conversely, structural divergence is proportional to mutational sensitivity [73]. Only mutations targeting critical regions such as rigid hinges could thus have the potential to change ENM mode patterns and function, causing either disease or driving evolution. The majority of changes would have no effect due to mode robustness, which would be the primary factor for evolutionary conservation. Apart from mode robustness, protein modularity and size also contribute to the overlap between the NMs and evolutionary modes and explain their low dimensionality, according to recent studies [111]. Altogether, these studies point out that biophysical properties are key for mode conservation.

Nevertheless, the functional motions observed experimentally seem to correlate with the soft modes more than expected based on just their amplitude and energies, indicating that selection plays a central role [112]. ENM studies indicate that selection guides sequence evolution to favor dynamical properties required for function, such as allosteric behavior or protein–protein interactions [113,114]. An exhaustive study by the Bahar group on nearly 27 K proteins representing 116 CATH superfamilies [72] characterized the cooperative mechanisms and convergent/divergent features that underlie the shared/differentiated dynamics of family members, developing an integrated pipeline to evaluate the signature dynamics of families based on ENMs (SignDy). They confirmed that global lowest-frequency modes of motion are conserved within a family, but there is a subset of motions that sharply distinguishes subfamilies at low-to-intermediate frequencies and is responsible for functional differentiation. Then, modulation of robust/conserved global dynamics via low-to-intermediate frequency fluctuations could be a versatile mechanism ensuring fold adaptability and subfamily specificity, subject to both positive and negative selection. Finally, taking one step further with this “selectionist” view, recent works have attempted to predict functional dynamics directly from sequence evolutionary couplings, skipping structures altogether [115].

4. Examples of Evolutionary Conservation, Convergence and Divergence

As we have seen, it is extremely difficult to disentangle the relevance of sequence, structure, and dynamics for evolutionary selection as they are intertwined. Database-wide comparative quantitative studies of protein dynamics are essential, but it is also important to keep in mind that, in the biological realm, “the devil can be in the details”, and a closer look at key conserved systems can be illuminating to understand how and to what extent evolution polishes protein shape and motions (Figure 1c,d). This is especially true for proteins executing the most fundamental life processes, prevalent in almost all living species; it is also true for the disease almost intrinsic to the mechanisms of pluricellular life, cancer, which can be viewed as an evolutionary process in miniature [116]. For example, it is well known that cells critically depend on pH and ion homeostasis, as well as membrane transport. Unsurprisingly, solute carriers and ion channels mediating these processes are incredibly well conserved from bacteria to humans, despite diverging 2–4 billion years ago [117,118]. Despite very low sequence identities, prokaryotic and eukaryotic versions of proteins such as cation/proton antiporters (CPAs), major facilitator superfamily transporters (MFSs), or pentameric ligand-gated ion channels (PLGICs), are incredibly conserved from a structural and conformational point of view. CPAs mediate the exchange of protons and monovalent cations such as Na^+ or K^+ , while MFS facilitates the movement of small solutes in response to gradients through cell membranes. Both MFSs and CPAs operate through an alternating-access mechanism, which requires a transition between states, where the substrate-binding site is exposed to opposite sides of the membrane alternately [119]. Structures show that MFSs follow a “rocker-switch” or “rocking bundle” mechanism, where the substrate-binding site is located at the interface of the so-called “transport” and “scaffold” domains. In contrast, CPAs work through an “elevator mechanism”, where the substrate-binding site is confined largely to a single “transport” domain that traverses the membrane along a relatively rigid, immobile, and central “core”. In the first, the barrier re-shapes and moves across the membrane while the substrate stays,

while in the second, it stays at a fixed position, and it is the substrate that moves across it. Both transport mechanisms are dependent on large-scale transitions between the so-called “inward” and “outward” states. Remarkably, despite sequence identities around just 20%, structures of the mammal SLC/NHE CPA family of Na^+/H^+ exchangers bear striking similarity with prokaryotic ones, like those of bacterial *Thermus thermophilus* NapA, archaeal *Pyrococcus abyssi* PaNhaP or *Methanocaldococcus jannaschii* MjNhaP1. This makes it possible to extract a highly conserved structural core (756 residues per homodimer) to achieve an incredibly low RMSD near $3.0 \pm 1.3 \text{ \AA}$ [74], which corresponds to the conformational transition tracked in the ensemble—when only one conformation is included, RMSD drops to 2 \AA , close to thermal fluctuations (Figure 2 and Table 1). Both bacterial and mammal structures are thus solved in inward- and outward-facing states, and therefore, their core ensemble’s main Principal Component (PC, see [26,60]) tracks the elevator motion responsible for transport. Significantly, this motion is also encoded in each one of the proteins: there is a high overlap (70–80%) between the transitions seen in the prokaryotic–eukaryotic ensemble and the lowest-frequency ENM modes from every individual member (Figure 2). Similarly, for MFSs, it is also possible to build a eukaryotic–prokaryotic “core” ensemble (353 residues) encompassing human, bovine, and rat GLUTs to *Plasmodium* PfHT1 or *Escherichia coli* Xyle [120], that despite the sequence identity around 30% has an RMSD as low as $2.7 \pm 1.2 \text{ \AA}$ and extremely similar rocking-bundle movements embedded on each structure. In the case of PLGICs, the notable resemblance between eukaryotic neurotransmitter channels and their simple prokaryotic counterparts like *Gloeobacter* GLIC has turned the latter into the perfect model to study gating mechanisms. As often happens with ancestral protein machines, their function (channel opening/closing) requires complex motions (extracellular blooming coupled to tilting/twisting of intracellular pore-gating helices), which are both embedded in their pentameric ring-like architecture and extremely conserved across evolution [55,121,122].

Table 1. Sequence, structural and dynamical similarity between mammalian NHE9 and bacterial proton exchanger NapA ¹.

	Identity	Similarity	TM-Score
<i>NHE9—NapA</i>	22%	42%	0.82
<i>Overlap NHE9—NapA NMA</i>	75%		
<i>Overlap NHE9—NapA X-ray transition</i>	82%		

¹ Adapted from Ref. [74].

Finally, another example of evolutionary selection acting on conformation could be behind mutational asymmetries in cancer, which tend to target signaling proteins. Global dynamics is a predictor of missense mutation pathogenicity [105,123] and in cancer genes, it has been shown that mutations tend to cluster in specific functional spots and specifically hinge regions as determined via ENMs [104]. One striking example is the oncogene EGFR, which displays a puzzling tissue-specific mutational asymmetry. In brain glioblastoma (GBM), mutations are highly heterogeneous but tend to cluster on the extracellular ligand-binding domain (ectodomain, ECD), even coexisting in the same tumor. In contrast, mutations in lung cancer concentrate in the intracellular kinase domain (KD), mostly focused on the catalytic cleft. This asymmetry results in intriguingly opposite responses to drugs binding to different KD conformers. Our ENM study of the ECD revealed that GBM mutations neatly cluster at hinge and interdomain regions, which control a large-scale conformational change of nearly 25 \AA between the closed-unbound and open-bound states. Further MD simulations revealed that GBM mutations favor spontaneous ECD opening following the lowest frequency modes, to acquire a transient conformation known to exist but never trapped experimentally. This ENM/MD intermediate was validated through structural, in vitro, and in vivo experiments [75,124,125], is shared by missense mutants from different ECD hotspots, and mimics the configuration of the most frequent change in GBM, the deletion EGFRvIII (Figure 1d). Specifically, the first tandem repeat of EGFR is

deleted in EGFRvIII but rotates in missense mutations. The ultimate goal of this remarkable structural “equivalence” or “convergence” trick is to allosterically activate the KD in a specific way, distinct from that favored by lung cancer mutations, which explains their different sensitivity to drugs. Importantly, lung and brain cancer mutations are known to activate different signaling pathways [126], and our ENM–MD studies suggest that this is directly governed by the different conformational dynamics they favor. On one hand, this could be an example of convergent evolution of missense mutations and deletions to achieve a similar functional outcome, driven by positive selection of those variants that explore the soft modes opening the structure in a “GBM-preferred” mode. On the other, the same protein, EGFR, apparently experiences divergent evolutionary trajectories in GBMs versus lung cancer to fine-tune its conformation and trigger cell growth in different niches—a potentially compelling case of evolution selecting lowest-frequency dynamics to modulate function.

In summary, the examples discussed above provide food for thought to question both the “selectionist-functional” view and the “biophysical-energetical” view of protein structure and dynamics evolution. Some works have focused on the interpretation of flexibility patterns under a predominantly evolutionary prism, while others favor the idea that the main cause of structural–dynamical divergence lies in the physical properties of proteins, such as their sensitivity to mutations. Observing the degree of conservation in ancestral proteins such as CPAs over scales of billions of years, despite having sequence identities in the “twilight” zone, strongly suggests a role for natural selection to keep key functional, structure-embedded mechanisms intact, especially for those proteins performing the most fundamental cellular tasks. These intrinsic motions have survived almost intact, from archaeobacteria to the human species, probably because of both their biophysical robustness and their biological fitness. Conversely, the striking clustering of mutations observed in cancer proteins to modulate not only their intrinsic dynamics but also their interactions with other proteins, etc., shows that, at high mutational rates and under selection pressure, evolution can quickly remodel and adapt what we could call protein “molecular phenotypes” [77], directly determined by their conformational dynamics and the resulting biological function. Importantly, there is mounting evidence that even local dynamics coupled to processes such as enzyme catalysis show clear footprints of evolutionary selection [127–131]. Looking forward, there are wide opportunities to apply ENMs to deepen studies of molecular evolution, which can illuminate its connections with protein biophysics or even guide protein design [132]. From analysis of the conservation of flexible versus rigid regions and how they relate to function, to evolutionarily classifying proteins based on their shape-encoded dynamics rather than strict sequence information, ENMs will allow us to explore the interplay of flexibility and evolutionary changes in the different kingdoms to an extent never imagined before, even more thanks to the incredibly expanded structural spaces that AI has opened [133,134].

Overall, we foresee that as experimental and computational evidence accumulates, and the increasingly active research on ENMs and evolution develops, we might reach a new paradigm. One in which biomolecular dynamics and, specifically, the large-scale motions intrinsic to 3D structures, could effectively be considered what biologist Ernst Mayr called “an object of selection” [135] at the most basic, microscopic scale of life.

Funding: This research was funded by Karolinska Institute, the Swedish Foundations for Cancer Research (Cancerfonden Junior Investigator Award CF 21 0305 JIA and Project Grant CF 21 1471 Pj), the Swedish Scientific Research Council (Vetenskapsrådet, VR 2021-02248) and the Jeansson, Hedlund and Sagen Foundations.

Data Availability Statement: Data used to generate the figures are available upon request.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Anfinsen, C.B.; Haber, E.; Sela, M.; White, F.H. The Kinetics of Formation of Native Ribonuclease during Oxidation of the Reduced Polypeptide Chain. *Proc. Natl. Acad. Sci. USA* **1961**, *47*, 1309–1314. [[CrossRef](#)] [[PubMed](#)]
2. Daggett, V.; Fersht, A. The Present View of the Mechanism of Protein Folding. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 497–502. [[CrossRef](#)] [[PubMed](#)]
3. Karplus, M.; McCammon, J.A. The Dynamics of Proteins. *Sci. Am.* **1986**, *254*, 42–51. [[CrossRef](#)] [[PubMed](#)]
4. Frank, J. New Opportunities Created by Single-Particle Cryo-EM: The Mapping of Conformational Space. *Biochemistry* **2018**, *57*, 888. [[CrossRef](#)] [[PubMed](#)]
5. James, L.C.; Tawfik, D.S. Conformational Diversity and Protein Evolution—A 60-Year-Old Hypothesis Revisited. *Trends Biochem. Sci.* **2003**, *28*, 361–368. [[CrossRef](#)] [[PubMed](#)]
6. Henzler-Wildman, K.; Kern, D. Dynamic Personalities of Proteins. *Nature* **2007**, *450*, 964–972. [[CrossRef](#)]
7. Karplus, M.; Kuriyan, J. Molecular Dynamics and Protein Function. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6679–6685. [[CrossRef](#)]
8. Karplus, M.; McCammon, J.A. Molecular Dynamics Simulations of Biomolecules. *Nat. Struct. Biol.* **2002**, *9*, 646–652. [[CrossRef](#)]
9. Amaral, M.; Kokh, D.B.; Bomke, J.; Wegener, A.; Buchstaller, H.P.; Eggenweiler, H.M.; Matias, P.; Sirrenberg, C.; Wade, R.C.; Frech, M. Protein Conformational Flexibility Modulates Kinetics and Thermodynamics of Drug Binding. *Nat. Commun.* **2017**, *8*, 2276. [[CrossRef](#)]
10. Tuffery, P.; Derreumaux, P. Flexibility and Binding Affinity in Protein–Ligand, Protein–Protein and Multi-Component Protein Interactions: Limitations of Current Computational Approaches. *J. R. Soc. Interface* **2012**, *9*, 20–33. [[CrossRef](#)]
11. Teague, S.J. Implications of Protein Flexibility for Drug Discovery. *Nat. Rev. Drug Discov.* **2003**, *2*, 527–541. [[CrossRef](#)] [[PubMed](#)]
12. Changeux, J.-P.; Edelstein, S. Conformational Selection or Induced-Fit? 50 Years of Debate Resolved. *F1000 Biol. Rep.* **2011**, *3*, 1–15. [[CrossRef](#)] [[PubMed](#)]
13. Csermely, P.; Palotai, R.; Nussinov, R. Induced Fit, Conformational Selection and Independent Dynamic Segments: An Extended View of Binding Events. *Trends Biochem. Sci.* **2010**, *35*, 539–546. [[CrossRef](#)]
14. Thulasingham, M.; Orellana, L.; Nji, E.; Ahmad, S.; Rinaldo-Matthis, A.; Haeggström, J.Z. Crystal Structures of Human MGST2 Reveal Synchronized Conformational Changes Regulating Catalysis. *Nat. Commun.* **2021**, *12*, 5721. [[CrossRef](#)]
15. Mhashal, A.R.; Romero-Rivera, A.; Mydy, L.S.; Cristobal, J.R.; Gulick, A.M.; Richard, J.P.; Kamerlin, S.C.L. Modeling the Role of a Flexible Loop and Active Site Side Chains in Hydride Transfer Catalyzed by Glycerol-3-Phosphate Dehydrogenase. *ACS Catal.* **2020**, *10*, 11253–11267. [[CrossRef](#)]
16. Henzler-Wildman, K.A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M.A.; Petsko, G.A.; Karplus, M.; et al. Intrinsic Motions along an Enzymatic Reaction Trajectory. *Nature* **2007**, *450*, 838–844. [[CrossRef](#)] [[PubMed](#)]
17. Babu, M.M.; Van Der Lee, R.; De Groot, N.S.; Gsponer, J. Intrinsically Disordered Proteins: Regulation and Disease. *Curr. Opin. Struct. Biol.* **2011**, *21*, 432–440. [[CrossRef](#)]
18. Uversky, V.N. Intrinsically Disordered Proteins and Their “Mysterious” (Meta)Physics. *Front. Phys.* **2019**, *7*, 10. [[CrossRef](#)]
19. Elber, R.; Karplus, M. Multiple Conformational States of Proteins: A Molecular Dynamics Analysis of Myoglobin. *Science* **1987**, *235*, 318–321. [[CrossRef](#)]
20. Bahar, I.; Lezon, T.R.; Yang, L.-W.; Eyal, E. Global Dynamics of Proteins: Bridging between Structure and Function. *Annu. Rev. Biophys.* **2010**, *39*, 23–42. [[CrossRef](#)]
21. Bastolla, U. Computing Protein Dynamics from Protein Structure with Elastic Network Models. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 488–503. [[CrossRef](#)]
22. López-Blanco, J.R.; Chacón, P. New Generation of Elastic Network Models. *Curr. Opin. Struct. Biol.* **2016**, *37*, 46–53. [[CrossRef](#)] [[PubMed](#)]
23. Sanejouand, Y.-H. Elastic Network Models: Theoretical and Empirical Foundations. *Network* **2011**, *26*, 601–616.
24. Bonomi, M.; Vendruscolo, M. Determination of Protein Structural Ensembles Using Cryo-Electron Microscopy. *Curr. Opin. Struct. Biol.* **2019**, *56*, 37–450. [[CrossRef](#)] [[PubMed](#)]
25. Krieger, J.M.; Sorzano, C.O.S.; Carazo, J.M.; Bahar, I. Protein Dynamics Developments for the Large Scale and CryoEM: Case Study of ProDy 2.0. *Acta Cryst. D Struct. Biol.* **2022**, *78*, 399–409. [[CrossRef](#)]
26. Orellana, L. Large-Scale Conformational Changes and Protein Function: Breaking the in Silico Barrier. *Front. Mol. Biosci.* **2019**, *6*, 117. [[CrossRef](#)]
27. Herzberg, G. *Molecular Spectra and Molecular Structure*; D. Van Nostrand Company, Inc.: Princeton, NJ, USA, 1945.
28. Wilson, E.B.; Decius, J.C.; Cross, P.C. *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra*; McGraw-Hill: New York, NY, USA, 1955.
29. Brooks, B. Harmonic Dynamics of Proteins: Normal Modes and Fluctuations in Bovine Pancreatic Trypsin Inhibitor. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 6571–6575. [[CrossRef](#)]
30. Go, N.; Noguti, T.; Nishikawa, T. Dynamics of a Small Globular Protein in Terms of Low-Frequency Vibrational Modes. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 3696–3700. [[CrossRef](#)]
31. Levitt, M.; Sander, C.; Stern, P.S. The normal modes of a protein: Native bovine pancreatic trypsin inhibitor. *Int. J. Quantum Chem.* **1983**, *24*, 181–199. [[CrossRef](#)]
32. Noguti, T.; Gō, N. Collective Variable Description of Small-Amplitude Conformational Fluctuations in a Globular Protein. *Nature* **1982**, *296*, 776–778. [[CrossRef](#)]

33. Tasumi, M.; Takeuchi, H.; Ataka, S.; Dwivedi, A.M.; Krimm, S. Normal Vibrations of Proteins: Glucagon. *Biopolymers* **1982**, *21*, 711–714. [[CrossRef](#)] [[PubMed](#)]
34. Orozco, M.; Orellana, L.; Hospital, A.; Naganathan, A.N.; Emperador, A.; Carrillo, O.; Gelpí, J.L. Coarse-Grained Representation of Protein Flexibility. Foundations, Successes, and Shortcomings. *Adv. Protein Chem. Struct. Biol.* **2011**, *85*, 183–215. [[CrossRef](#)] [[PubMed](#)]
35. Flory, P.J.; Gordon, M.; McCrum, N.G. Statistical Thermodynamics of Random Networks [and Discussion]. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1976**, *351*, 351–380. [[CrossRef](#)]
36. Rouse, P.E. A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers. *J. Chem. Phys.* **1953**, *21*, 1272. [[CrossRef](#)]
37. Tirion, M. Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. *Phys. Rev. Lett.* **1996**, *77*, 1905–1908. [[CrossRef](#)]
38. Bahar, I.; Atilgan, A.R.; Erman, B. Direct Evaluation of Thermal Fluctuations in Proteins Using a Single-Parameter Harmonic Potential. *Fold. Des.* **1997**, *2*, 173–181. [[CrossRef](#)]
39. Atilgan, A.R.; Durell, S.R.; Jernigan, R.L.; Demirel, M.C.; Keskin, O.; Bahar, I. Anisotropy of Fluctuation Dynamics of Proteins with an Elastic Network Model. *Biophys. J.* **2001**, *80*, 505–515. [[CrossRef](#)]
40. Bauer, J.A.; Pavlović, J.; Bauerová-Hlinková, V. Normal Mode Analysis as a Routine Part of a Structural Investigation. *Molecules* **2019**, *24*, 3293. [[CrossRef](#)]
41. Dobbins, S.E.; Lesk, V.I.; Sternberg, M.J.E. Insights into Protein Flexibility: The Relationship between Normal Modes and Conformational Change upon Protein-Protein Docking. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10390–10395. [[CrossRef](#)]
42. Petrone, P.; Pande, V.S. Can Conformational Change Be Described by Only a Few Normal Modes? *Biophys. J.* **2006**, *90*, 1583–1593. [[CrossRef](#)]
43. Stein, A.; Rueda, M.; Panjkovich, A.; Orozco, M.; Aloy, P. A Systematic Study of the Energetics Involved in Structural Changes upon Association and Connectivity in Protein Interaction Networks. *Structure* **2011**, *19*, 881–889. [[CrossRef](#)] [[PubMed](#)]
44. Yang, L.; Song, G.; Jernigan, R.L. How Well Can We Understand Large-Scale Protein Motions Using Normal Modes of Elastic Network Models? *Biophys. J.* **2007**, *93*, 920–929. [[CrossRef](#)] [[PubMed](#)]
45. Vu, H.T.; Zhang, Z.; Tehver, R.; Thirumalai, D. Plus and Minus Ends of Microtubules Respond Asymmetrically to Kinesin Binding by a Long-Range Directionally Driven Allosteric Mechanism. *Sci. Adv.* **2022**, *8*, eabn0856. [[CrossRef](#)] [[PubMed](#)]
46. Kolan, D.; Fonar, G.; Samson, A.O. Elastic Network Normal Mode Dynamics Reveal the GPCR Activation Mechanism. *Proteins Struct. Funct. Bioinform.* **2014**, *82*, 579–586. [[CrossRef](#)]
47. Bahar, I. On the Functional Significance of Soft Modes Predicted by Coarse-Grained Models for Membrane Proteins. *J. Gen. Physiol.* **2010**, *135*, 563–573. [[CrossRef](#)]
48. Isin, B.; Rader, A.J.; Dhiman, H.K.; Klein-Seetharaman, J.; Bahar, I. Predisposition of the Dark State of Rhodopsin to Functional Changes in Structure. *Proteins Struct. Funct. Bioinform.* **2006**, *65*, 970–983. [[CrossRef](#)]
49. Gerstein, M.; Krebs, W. A Database of Macromolecular Motions. *Nucleic Acids Res.* **1998**, *26*, 4280–4290. [[CrossRef](#)]
50. Krebs, W.G.; Alexandrov, V.; Wilson, C.A.; Echols, N.; Yu, H.; Gerstein, M. Normal Mode Analysis of Macromolecular Motions in a Database Framework: Developing Mode Concentration as a Useful Classifying Statistic. *Proteins* **2002**, *48*, 682–695. [[CrossRef](#)]
51. Tama, F.; Sanejouand, Y.H. Conformational Change of Proteins Arising from Normal Mode Calculations. *Protein Eng.* **2001**, *14*, 1–6. [[CrossRef](#)]
52. Alexandrov, V. Normal Modes for Predicting Protein Motions: A Comprehensive Database Assessment and Associated Web Tool. *Protein Sci.* **2005**, *14*, 633–643. [[CrossRef](#)]
53. Gur, M.; Zomot, E.; Bahar, I. Global Motions Exhibited by Proteins in Micro- to Milliseconds Simulations Concur with Anisotropic Network Model Predictions. *J. Chem. Phys.* **2013**, *139*, 121912. [[CrossRef](#)]
54. Orellana, L.; Rueda, M.; Ferrer-Costa, C.; Lopez-Blanco, J.R.; Chacón, P.; Orozco, M. Approaching Elastic Network Models to Molecular Dynamics Flexibility. *J. Chem. Theory Comput.* **2010**, *6*, 2910–2923. [[CrossRef](#)] [[PubMed](#)]
55. Orellana, L.; Yoluk, O.; Carrillo, O.; Orozco, M.; Lindahl, E. Prediction and Validation of Protein Intermediate States from Structurally Rich Ensembles and Coarse-Grained Simulations. *Nat. Commun.* **2016**, *7*, 12575. [[CrossRef](#)] [[PubMed](#)]
56. Yang, L.; Song, G.; Carriquiry, A.; Jernigan, R.L. Close Correspondence between the Motions from Principal Component Analysis of Multiple HIV-1 Protease Structures and Elastic Network Modes. *Structure* **2008**, *16*, 321–330. [[CrossRef](#)] [[PubMed](#)]
57. Rueda, M.; Chacón, P.; Orozco, M. Thorough Validation of Protein Normal Mode Analysis: A Comparative Study with Essential Dynamics. *Structure* **2007**, *15*, 565–575. [[CrossRef](#)]
58. Daidone, I.; Amadei, A. Essential Dynamics: Foundation and Applications. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 762–770. [[CrossRef](#)]
59. Jolliffe, I.T.; Cadima, J. Principal Component Analysis: A Review and Recent Developments. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2016**, *374*, 20150202. [[CrossRef](#)]
60. Kitao, A. Principal Component Analysis and Related Methods for Investigating the Dynamics of Biological Macromolecules. *J* **2022**, *5*, 298–317. [[CrossRef](#)]
61. Rueda, M.; Ferrer-Costa, C.; Meyer, T.; Pérez, A.; Camps, J.; Hospital, A.; Gelpí, J.L.; Orozco, M. A Consensus View of Protein Dynamics. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 796–801. [[CrossRef](#)]

62. Meyer, T.; D'Abramo, M.; Hospital, A.; Rueda, M.; Ferrer-Costa, C.; Pérez, A.; Carrillo, O.; Camps, J.; Fenollosa, C.; Repchevsky, D.; et al. MoDEL (Molecular Dynamics Extended Library): A Database of Atomistic Molecular Dynamics Trajectories. *Structure* **2010**, *18*, 1399–1409. [[CrossRef](#)]
63. Ma, J. Usefulness and Limitations of Normal Mode Analysis in Modeling Dynamics of Biomolecular Complexes. *Structure* **2005**, *13*, 373–380. [[CrossRef](#)] [[PubMed](#)]
64. Hubbard, J.H. The KAM Theorem. In *Kolmogorov's Heritage in Mathematics*; Charpentier, É., Lesne, A., Nikolski, N.K., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; pp. 215–238. ISBN 978-3-540-36351-4.
65. Kern, D.; Zuiderweg, E.R. The Role of Dynamics in Allosteric Regulation. *Curr. Opin. Struct. Biol.* **2003**, *13*, 748–757. [[CrossRef](#)] [[PubMed](#)]
66. Goh, C.-S.; Milburn, D.; Gerstein, M. Conformational Changes Associated with Protein-Protein Interactions. *Curr. Opin. Struct. Biol.* **2004**, *14*, 104–109. [[CrossRef](#)]
67. Hayward, S.; Kitao, A.; Go, N. Harmonic and Anharmonic Aspects in the Dynamics of BPTI: A Normal Mode Analysis and Principal Component Analysis. *Protein Sci. A Publ. Protein Soc.* **1994**, *3*, 936–943. [[CrossRef](#)] [[PubMed](#)]
68. Hayward, S.; Kitao, A.; Go, N. Harmonicity and Anharmonicity in Protein Dynamics: A Normal Mode Analysis and Principal Component Analysis. *Proteins* **1995**, *23*, 177–186. [[CrossRef](#)]
69. Dehouck, Y.; Bastolla, U. Why Are Large Conformational Changes Well Described by Harmonic Normal Modes? *Biophys. J.* **2021**, *120*, 5343–5354. [[CrossRef](#)]
70. Kaynak, B.T.; Krieger, J.M.; Dudas, B.; Dahmani, Z.L.; Costa, M.G.S.; Balog, E.; Scott, A.L.; Doruker, P.; Perahia, D.; Bahar, I. Sampling of Protein Conformational Space Using Hybrid Simulations: A Critical Assessment of Recent Methods. *Front. Mol. Biosci.* **2022**, *9*, 832847. [[CrossRef](#)]
71. Amadei, A.; Linssen, A.B.; Berendsen, H.J. Essential Dynamics of Proteins. *Proteins* **1993**, *17*, 412–425. [[CrossRef](#)] [[PubMed](#)]
72. Zhang, S.; Li, H.; Krieger, J.M.; Bahar, I. Shared Signature Dynamics Tempered by Local Fluctuations Enables Fold Adaptability and Specificity. *Mol. Biol. Evol.* **2019**, *36*, 2053–2068. [[CrossRef](#)]
73. Marcos, M.L.; Echave, J. The Variation among Sites of Protein Structure Divergence Is Shaped by Mutation and Scaled by Selection. *Curr. Res. Struct. Biol.* **2020**, *2*, 156–163. [[CrossRef](#)]
74. Winkelmann, I.; Matsuoka, R.; Meier, P.F.; Shutin, D.; Zhang, C.; Orellana, L.; Sexton, R.; Landreh, M.; Robinson, C.V.; Beckstein, O.; et al. Structure and Elevator Mechanism of the Mammalian Sodium/Proton Exchanger NHE9. *EMBO J.* **2020**, *39*, 4541–4559. [[CrossRef](#)] [[PubMed](#)]
75. Orellana, L.; Thorne, A.H.; Lema, R.; Gustavsson, J.; Parisian, A.D.; Hospital, A.; Cordeiro, T.N.; Bernadó, P.; Scott, A.M.; Brun-Heath, I.; et al. Oncogenic Mutations at the EGFR Ectodomain Structurally Converge to Remove a Steric Hindrance on a Kinase-Coupled Cryptic Epitope. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 10009–10018. [[CrossRef](#)]
76. Muñoz, M.M.; Price, S.A. The Future Is Bright for Evolutionary Morphology and Biomechanics in the Era of Big Data. *Integr. Comp. Biol.* **2019**, *59*, 599–603. [[CrossRef](#)] [[PubMed](#)]
77. Sikosek, T.; Chan, H.S. Biophysics of Protein Evolution and Evolutionary Protein Biophysics. *J. R. Soc. Interface* **2014**, *11*, 20140419. [[CrossRef](#)] [[PubMed](#)]
78. Fuglebakk, E.; Tiwari, S.P.; Reuter, N. Comparing the Intrinsic Dynamics of Multiple Protein Structures Using Elastic Network Models. *Biochim. Biophys. Acta—Gen. Subj.* **2015**, *1850*, 911–922. [[CrossRef](#)] [[PubMed](#)]
79. Bordin, N.; Sillitoe, I.; Lees, J.G.; Orenco, C. Tracing Evolution Through Protein Structures: Nature Captured in a Few Thousand Folds. *Front. Mol. Biosci.* **2021**, *8*, 668184. [[CrossRef](#)]
80. Chothia, C.; Lesk, A.M. The Relation between the Divergence of Sequence and Structure in Proteins. *EMBO J.* **1986**, *5*, 823–826. [[CrossRef](#)]
81. Kuhlman, B.; Baker, D. Native Protein Sequences Are Close to Optimal for Their Structures. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 10383–10388. [[CrossRef](#)]
82. Koehl, P.; Levitt, M. Protein Topology and Stability Define the Space of Allowed Sequences. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1280–1285. [[CrossRef](#)]
83. Zheng, W.; Brooks, B.R.; Thirumalai, D. Low-Frequency Normal Modes That Describe Allosteric Transitions in Biological Nanomachines Are Robust to Sequence Variations. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 7664–7669. [[CrossRef](#)]
84. Hensen, U.; Meyer, T.; Haas, J.; Rex, R.; Vriend, G.; Grubmüller, H. Exploring Protein Dynamics Space: The Dynasome as the Missing Link between Protein Structure and Function. *PLoS ONE* **2012**, *7*, e33931. [[CrossRef](#)] [[PubMed](#)]
85. Narunsky, A.; Nepomnyachiy, S.; Ashkenazy, H.; Kolodny, R.; Ben-Tal, N. ConTemplate Suggests Possible Alternative Conformations for a Query Protein of Known Structure. *Structure* **2015**, *23*, 2162–2170. [[CrossRef](#)] [[PubMed](#)]
86. Zsolanyi, F.; Ambrus, V.; Fuxreiter, M. Patterns of Dynamics Comprise a Conserved Evolutionary Trait. *J. Mol. Biol.* **2020**, *432*, 497–507. [[CrossRef](#)] [[PubMed](#)]
87. Bastolla, U.; Dehouck, Y.; Echave, J. What Evolution Tells Us about Protein Physics, and Protein Physics Tells Us about Evolution. *Curr. Opin. Struct. Biol.* **2017**, *42*, 59–66. [[CrossRef](#)] [[PubMed](#)]
88. Liberles, D.A.; Teichmann, S.A.; Bahar, I.; Bastolla, U.; Bloom, J.; Bornberg-Bauer, E.; Colwell, L.J.; De Koning, A.P.J.; Dokholyan, N.V.; Echave, J.; et al. The Interface of Protein Structure, Protein Biophysics, and Molecular Evolution. *Protein Sci.* **2012**, *21*, 769–785. [[CrossRef](#)] [[PubMed](#)]

89. Tiwari, S.P.; Reuter, N. Conservation of Intrinsic Dynamics in Proteins—What Have Computational Models Taught Us? *Curr. Opin. Struct. Biol.* **2018**, *50*, 75–81. [[CrossRef](#)] [[PubMed](#)]
90. Keskin, O.; Jernigan, R.L.; Bahar, I. Proteins with Similar Architecture Exhibit Similar Large-Scale Dynamic Behavior. *Biophys. J.* **2000**, *78*, 2093–2106. [[CrossRef](#)]
91. Leo-Macias, A.; Lopez-Romero, P.; Lupyán, D.; Zerbino, D.; Ortiz, A.R. An Analysis of Core Deformations in Protein Superfamilies. *Biophys. J.* **2005**, *88*, 1291–1299. [[CrossRef](#)]
92. Velázquez-Muriel, J.A.; Rueda, M.; Cuesta, I.; Pascual-Montano, A.; Orozco, M.; Carazo, J.-M. Comparison of Molecular Dynamics and Superfamily Spaces of Protein Domain Deformation. *BMC Struct. Biol.* **2009**, *9*, 6. [[CrossRef](#)]
93. Leo-Macias, A.; Lopez-Romero, P.; Lupyán, D.; Zerbino, D.; Ortiz, A.R. Core Deformations in Protein Families: A Physical Perspective. *Biophys. Chem.* **2005**, *115*, 125–128. [[CrossRef](#)]
94. Maguid, S.; Fernández-Alberti, S.; Ferrelli, L.; Echave, J. Exploring the Common Dynamics of Homologous Proteins. Application to the Globin Family. *Biophys. J.* **2005**, *89*, 3–13. [[CrossRef](#)] [[PubMed](#)]
95. Maguid, S.; Fernández-Alberti, S.; Parisi, G.; Echave, J. Evolutionary Conservation of Protein Backbone Flexibility. *J. Mol. Evol.* **2006**, *63*, 448–457. [[CrossRef](#)] [[PubMed](#)]
96. Skjaerven, L.; Yao, X.Q.; Scarabelli, G.; Grant, B.J. Integrating Protein Structural Dynamics and Evolutionary Analysis with Bio3D. *BMC Bioinform.* **2014**, *15*, 399. [[CrossRef](#)]
97. Franzosa, E.A.; Xia, Y. Structural Determinants of Protein Evolution Are Context-Sensitive at the Residue Level. *Mol. Biol. Evol.* **2009**, *26*, 2387–2395. [[CrossRef](#)] [[PubMed](#)]
98. Huang, T.-T.; del Valle Marcos, M.L.; Hwang, J.-K.; Echave, J. A Mechanistic Stress Model of Protein Evolution Accounts for Site-Specific Evolutionary Rates and Their Relationship with Packing Density and Flexibility. *BMC Evol. Biol.* **2014**, *14*, 78. [[CrossRef](#)]
99. Marsh, J.A.; Teichmann, S.A. Parallel Dynamics and Evolution: Protein Conformational Fluctuations and Assembly Reflect Evolutionary Changes in Sequence and Structure. *BioEssays* **2014**, *36*, 209–218. [[CrossRef](#)]
100. Dong, Z.; Zhou, H.; Tao, P. Combining Protein Sequence, Structure, and Dynamics: A Novel Approach for Functional Evolution Analysis of PAS Domain Superfamily. *Protein Sci.* **2018**, *27*, 421–430. [[CrossRef](#)]
101. Liu, Y.; Bahar, I. Sequence Evolution Correlates with Structural Dynamics. *Mol. Biol. Evol.* **2012**, *29*, 2253–2263. [[CrossRef](#)]
102. Campitelli, P.; Modi, T.; Kumar, S.; Ozkan, S.B. The Role of Conformational Dynamics and Allostery in Modulating Protein Evolution. *Annu. Rev. Biophys.* **2020**, *49*, 267–288. [[CrossRef](#)]
103. Nevin Gerek, Z.; Kumar, S.; Banu Ozkan, S. Structural Dynamics Flexibility Informs Function and Evolution at a Proteome Scale. *Evol. Appl.* **2013**, *6*, 423–433. [[CrossRef](#)]
104. Sayılğan, J.F.; Haliloğlu, T.; Gönen, M. Protein Dynamics Analysis Reveals That Missense Mutations in Cancer-Related Genes Appear Frequently on Hinge-Neighboring Residues. *Proteins Struct. Funct. Bioinform.* **2019**, *87*, 512–519. [[CrossRef](#)] [[PubMed](#)]
105. Ponzoni, L.; Bahar, I. Structural Dynamics Is a Determinant of the Functional Significance of Missense Variants. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4164–4169. [[CrossRef](#)]
106. Frappier, V.; Najmanovich, R.J. A Coarse-Grained Elastic Network Atom Contact Model and Its Use in the Simulation of Protein Dynamics and the Prediction of the Effect of Mutations. *PLoS Comput. Biol.* **2014**, *10*, e1003569. [[CrossRef](#)]
107. Banerjee, A.; Bahar, I. Structural Dynamics Predominantly Determine the Adaptability of Proteins to Amino Acid Deletions. *Int. J. Mol. Sci.* **2023**, *24*, 8450. [[CrossRef](#)]
108. Echave, J. Evolutionary Divergence of Protein Structure: The Linearly Forced Elastic Network Model. *Chem. Phys. Lett.* **2008**, *457*, 413–416. [[CrossRef](#)]
109. Echave, J.; Fernández, F.M. A Perturbative View of Protein Structural Variation. *Proteins Struct. Funct. Bioinform.* **2010**, *78*, 173–180. [[CrossRef](#)]
110. Echave, J. Why Are the Low-Energy Protein Normal Modes Evolutionarily Conserved? *Pure Appl. Chem.* **2012**, *84*, 1931–1937. [[CrossRef](#)]
111. Tang, Q.-Y.; Kaneko, K. Dynamics-Evolution Correspondence in Protein Structures. *Phys. Rev. Lett.* **2021**, *127*, 098103. [[CrossRef](#)]
112. Dos Santos, H.G.; Klett, J.; Méndez, R.; Bastolla, U. Characterizing Conformation Changes in Proteins through the Torsional Elastic Response. *Biochim. Biophys. Acta* **2013**, *1834*, 836–846. [[CrossRef](#)]
113. Haliloglu, T.; Bahar, I. Adaptability of Protein Structures to Enable Functional Interactions and Evolutionary Implications. *Curr. Opin. Struct. Biol.* **2015**, *35*, 17–23. [[CrossRef](#)] [[PubMed](#)]
114. Zhang, Y.; Doruker, P.; Kaynak, B.; Zhang, S.; Krieger, J.; Li, H.; Bahar, I. Intrinsic Dynamics Is Evolutionarily Optimized to Enable Allosteric Behavior. *Curr. Opin. Struct. Biol.* **2020**, *62*, 14–21. [[CrossRef](#)] [[PubMed](#)]
115. Jia, K.; Kilinc, M.; Jernigan, R.L. Functional Protein Dynamics Directly from Sequences. *J. Phys. Chem. B* **2023**, *127*, 1914–1921. [[CrossRef](#)] [[PubMed](#)]
116. Nowell, P.C. The Clonal Evolution of Tumor Cell Populations. *Science* **1976**, *194*, 23–28. [[CrossRef](#)]
117. Hedges, S.B.; Chen, H.; Kumar, S.; Wang, D.Y.; Thompson, A.S.; Watanabe, H. A Genomic Timescale for the Origin of Eukaryotes. *BMC Evol. Biol.* **2001**, *1*, 4. [[CrossRef](#)] [[PubMed](#)]
118. Long, X.; Xue, H.; Wong, J.T.-F. Descent of Bacteria and Eukarya From an Archaeal Root of Life. *Evol. Bioinform. Online* **2020**, *16*, 1176934320908267. [[CrossRef](#)]

119. Drew, D.; Boudker, O. Shared Molecular Mechanisms of Membrane Transporters. *Annu. Rev. Biochem.* **2016**, *85*, 543–572. [[CrossRef](#)] [[PubMed](#)]
120. Qureshi, A.A.; Suades, A.; Matsuoka, R.; Brock, J.; McComas, S.E.; Nji, E.; Orellana, L.; Claesson, M.; Delemotte, L.; Drew, D. The Molecular Basis for Sugar Import in Malaria Parasites. *Nature* **2020**, *578*, 321–325. [[CrossRef](#)]
121. Howard, R.J. Elephants in the Dark: Insights and Incongruities in Pentameric Ligand-Gated Ion Channel Models. *J. Mol. Biol.* **2021**, *433*, 167128. [[CrossRef](#)]
122. Mhashal, A.R.; Yoluk, O.; Orellana, L. Exploring the Conformational Impact of Novel Glycine Receptor Mutations through Coarse-Grained Analysis and Atomistic Simulations. *Front. Mol. Biosci.* **2022**, *9*, 890851. [[CrossRef](#)]
123. Ponzoni, L.; Peñaherrera, D.A.; Oltvai, Z.N.; Bahar, I. Rhapsody: Predicting the Pathogenicity of Human Missense Variants. *Bioinformatics* **2020**, *36*, 3084–3092. [[CrossRef](#)]
124. Orellana, L. Convergence of EGFR Glioblastoma Mutations: Evolution and Allosteric Rationalizing Targeted Therapy. *Mol. Cell. Oncol.* **2019**, *6*, e1630798. [[CrossRef](#)] [[PubMed](#)]
125. Orellana, L.; Hospital, A.; Orozco, M. Oncogenic Mutations of the EGF-Receptor Ectodomain Reveal an Unexpected Mechanism for Ligand-Independent Activation. *bioRxiv* **2014**. [[CrossRef](#)]
126. Uribe, M.L.; Marrocco, I.; Yarden, Y. EGFR in Cancer: Signaling Mechanisms, Drugs, and Acquired Resistance. *Cancers* **2021**, *13*, 2748. [[CrossRef](#)] [[PubMed](#)]
127. Lai, J.; Jin, J.; Kubelka, J.; Liberles, D.A. A Phylogenetic Analysis of Normal Modes Evolution in Enzymes and Its Relationship to Enzyme Function. *J. Mol. Biol.* **2012**, *422*, 442–459. [[CrossRef](#)]
128. Petrovic, D.; Risso, V.A.; Kamerlin, S.C.L.; Sanchez-Ruiz, J.M. Conformational Dynamics and Enzyme Evolution. *J. R. Soc. Interface* **2018**, *15*, 20180330. [[CrossRef](#)]
129. Narayanan, C.; Bernard, D.N.; Bafna, K.; Gagné, D.; Chennubhotla, C.S.; Doucet, N.; Agarwal, P.K. Conservation of Dynamics Associated with Biological Function in an Enzyme Superfamily. *Structure* **2018**, *26*, 426–436.e3. [[CrossRef](#)]
130. Ramanathan, A.; Agarwal, P.K. Evolutionarily Conserved Linkage between Enzyme Fold, Flexibility, and Catalysis. *PLoS Biol.* **2011**, *9*, e1001193. [[CrossRef](#)]
131. Carnevale, V.; Raugei, S.; Micheletti, C.; Carloni, P. Convergent Dynamics in the Protease Enzymatic Superfamily. *J. Am. Chem. Soc.* **2006**, *128*, 9766–9772. [[CrossRef](#)]
132. Campbell, E.C.; Correy, G.J.; Mabbitt, P.D.; Buckle, A.M.; Tokuriki, N.; Jackson, C.J. Laboratory Evolution of Protein Conformational Dynamics. *Curr. Opin. Struct. Biol.* **2018**, *50*, 49–57. [[CrossRef](#)]
133. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)]
134. Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; et al. Evolutionary-Scale Prediction of Atomic-Level Protein Structure with a Language Model. *Science* **2023**, *379*, 1123–1130. [[CrossRef](#)] [[PubMed](#)]
135. Mayr, E. The Objects of Selection. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 2091–2094. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.