In recent years, there has been a surge in the number of studies exploring the relationship between proteins’ equilibrium dynamics and structural changes involved in function. An emerging concept, supported by both theory and experiments, is that under native state conditions proteins have an intrinsic ability to sample conformations that meet functional requirements. A typical example is the ability of enzymes to sample open and closed forms, irrespective of substrate, succeeded by the stabilization of one form (usually closed) upon substrate binding. This ability is structure-encoded, and plays a key role in facilitating allosteric regulation, which suggests complementing the sequence-encodes-structure paradigm of protein science by structure-encodes-dynamics-encodes-function. The emerging connection implies an evolutionary role in selecting/conserving structures based on their ability to achieve functional dynamics, and in turn, selecting sequences that fold into such ‘apt’ structures.

Addresses
Department of Computational Biology, School of Medicine, University of Pittsburgh, Suite 3064, Biomedical Science Tower 3, 3051 Fifth Avenue, Pittsburgh, PA 15213, United States

Corresponding author: Bahar, Ivet (bahar@pitt.edu)

Intrinsic dynamics of enzymes in the unbound state and relation to allosteric regulation
Ivet Bahar, Chakra Chennubhotla and Dror Tobi

Introduction
Proteins sample an ensemble of conformations under equilibrium conditions. This ensemble is broadly distributed in the denatured state, and it becomes narrowly distributed – mainly confined to the neighborhood of the folded state – under native state conditions. Of interest are those conformations accessible near the global energy minimum, also called substates when separated by low energy barriers. Because proteins perform their function under these conditions, interconversions between these conformations are potentially functional.

The fact that folded proteins are not static, but undergo ‘wigglings and jigglings’ as put forth by Feynman, is now well-established. We have indeed come a long way since the idea was first put forward in the pioneering molecular dynamics (MD) simulations of proteins by McCammon, Karplus, Wolynes, Levitt, van Gunsteren and others in the late 1970s and early 1980s. With recent advances in experiments and theory, increased evidence is now being provided for the biological functionality of these apparently random motions. Experiments now permit us to visualize the structural flexibility and heterogeneity of biomolecules and assess their relevance to catalysis or signaling [1*,2**]. On the theoretical side, novel coarse-grained models and methods are providing insights into structure–dynamics relations on a global, rather than local, scale [3,4*]. The rapidly accumulating data lead to the emergence of concepts such as the pre-disposition or intrinsic ability of proteins to undergo conformational changes required for function, and a possible evolutionary pressure for selecting such structures, while also raising new questions with regard to old concepts.

New questions on old concepts
The first question concerns the conformational changes observed between the substrate-bound and substrate-unbound forms of a given enzyme, a phenomenon broadly referred to as ‘induced fit’ after the original proposition of Koshland [5]. The question is, to what extent these conformational changes are literally ‘induced’ by substrate. Would it be possible for the substrate to drive a change if the structure was not pre-disposed to undergo the change? Does the substrate simply stabilize the fittest conformations that already exist in the unbound state, following the redistribution of a pre-existing population [6] originally proposed by Weber [7]? Does it essentially select the lowest-energy-cost pathways away from the original minimum to optimize its interaction with the enzyme?

The second question relates to allosteric changes in conformations usually occurring on a large scale (quaternary changes) stated to be ‘driven’ by local phenomena like ATP binding or hydrolysis, ligand binding, phosphorylation, among others. Again, would it be possible to elicit cooperative responses, if these were not already energetically favored by the structure? How similar are the experimentally known allosteric changes, and those theoretically predicted to be naturally sampled by the particular structure?

Third, how can we reconcile the ensemble of conformational changes near folded state and the two-state transition observed in many allosteric proteins in accord...
Experimental evidence for conformational diversity of folded proteins and comparison with theoretical predictions. (a) Three conformations of cyclin dependent kinases (CDKs) adopted in the free form (middle), and in the presence of two different substrates, an inhibitor (INK4; left) and its activator (cyclin; right). The corresponding Protein data bank (PDB) codes are 1bi7, 1hcl and 1fin, in reading order. Colors refer to N-lobe (purple), C-lobe (red), hinge residues (orange) and activation loop (cyan). Both activation and inhibition involve conformational changes in and around the catalytic cleft. The activation loop (cyan) rotates towards the substrate (not shown). (b) Alternative conformations of HIV-1 RT. RT is composed of two subunits, p66 and p51 (wheat); the p66 subunit consists of two domains, polymerase and RNase H (blue); and the polymerase domain contains four subdomains, thumb (red), fingers (blue), palm (pink), connection (green). Comparison of the inhibitor-bound (nevirapine; space filling...
with the Monod–Wyman–Changeux (MWC) model [8]? Does all-or-none transition reflect passages between the most probable substates? Are there intermediates or sequence of events not detectable within experimental time frames? Do simulations reveal such features for small allosteric proteins or signaling proteins, which lack the regularity and symmetry that enhances co-operativity in multimeric quaternary structures? Given that certain key sites and interactions modulate the propagation of signals in many proteins, can we think of a structure-encoded network of interactions, and thereby a sequence of events, reminiscent of the Koshland–Némethy–Filmer (KNF) model [9]? Here, we will focus on the nature and functional significance of proteins’ intrinsic dynamics. First, we present recent data that point to the pre-existence, or accessibility, of functional conformations even in the absence of activity or triggering events. Second, we emphasize the fact that these motions are not random, but uniquely defined by the three-dimensional structure. In a sense, proteins appear to have optimally evolved to achieve their functional dynamics. Third, we call attention to the ability of structures to define, not only mechanisms of domain movements, to intrinsic disorder on a local scale. This fact that these motions are not random, but uniquely defined by the three-dimensional structure. In a sense, proteins appear to have optimally evolved to achieve their functional dynamics. Third, we call attention to the ability of structures to define, not only mechanisms of concerted motions, but also pathways of communication that ensure rapid propagation of perturbations and allosteric responses.

Intrinsic dynamics of proteins

Conformational variability in the presence of different substrates

Structural changes occur at multiple levels, ranging from concerted rearrangements of intact subunits or domain movements, to intrinsic disorder on a local scale. This ability to assume a well-defined ensemble of substates, near the ‘folded’ state indeed emerges as a consequence of the internal degrees of freedom that permit the structure to relax/rearrange without altering the fold. Figure 1a and b illustrates different conformations assumed by two well-studied enzymes in the presence of different substrates, (a) cyclin-dependent kinases (CDKs), and (b) HIV-1 reverse transcriptase (RT). CDK exhibits changes at the active loop conformation (cyan) and its N-lobe reorients relative to the C-lobe; RT undergoes global changes (e.g. movement of the thumb subdomain (red) on p66 subunit).

Such conformational variability observed in substrate-bound or ligand-bound forms is not restricted to enzymes [6], as illustrated in the seminal work of Frauenfelder and coworkers for myoglobin [10]. A recent significant study shows that the antibody SPE7 binds a range of antigens upon adopting different binding-site conformations in the free state [11*]. The crystal structure of the encounter complex on ligand-binding pathway reveals that the antibody becomes selective after a ‘postbinding’ conformational switch that stabilizes specific complexes, while others are dissociated [12**]. This study thus unravels a functional combination of prebinding equilibrium and postbinding (induced) fit that provides a discriminatory mechanism in ligand recognition [12**].

Accessibility of functional motions in the absence of substrate

While structural changes assumed in the presence of different substrates are common in proteins, what is new is their accessibility in the absence substrate. Also, motions involved in biophysical (e.g. substrate recognition and binding), biochemical (e.g. catalysis) or biological (e.g. signaling, transcription) activities are observed to be sampled even before activity.

The linkage between pre-existing dynamics and catalytic function is highlighted in recent reviews [1*13]. Motions characteristic of catalytic function were detected by Kern lab, for example, in the free states of prolyl cis–trans isomerase cyclophilin A, CypA [14*], and peptidylprolyl isomerase (Pin1) catalytic domain [15], leading to the conclusion that particular catalytic residues were ‘primed for catalysis’ [15]. Recent small angle X-ray scattering measurements also show that the D236A mutant of ATCase almost equally populates the T and R states [16], supporting the view of accessibility of alternative functional forms, detectable upon perturbing the structure at a critical site. These observations conform to the MWC model, where the R–T equilibrium is stated to be ‘an intrinsic property of allosteric oligomers accessible in the absence of ligand, with the ligand stabilizing the conformation to which it binds with higher affinity’ [2**].

(Figure 1 Legend Continued) form in yellow; left, unliganded (middle) and DNA-bound (right) forms a (PDB codes 1rh, 1dlo and 2ml) shows domain movements. (c) The left two diagrams display the free (left) and substrate-bound (middle) forms of adenylyl kinase (AK) (respective PDB codes: 4ake and 1ake), and the diagram on the right is a reconfigured form of the free enzyme computed by deforming the unbound structure along the lowest frequency ANM mode (mode 1) [18]. AK contains three domains: core (white), lid (cyan), and AMP-binding (green) domains. The substrate is shown in orange, space-filling representation. In the substrate-bound form as well as the model on the right, the lid approaches the core. (d) Comparison of the T (tense, unliganded) (left) and R2 (relaxed, CO-bound) (middle) forms of hemoglobin (Hb) (respectively PDB codes: 1a3n and 1bbb), and the model (right) calculated [19] by deforming the T form along ANM mode 2. The model approximates the experimentally observed torsion of the α2β2 dimer (front, colored) with respect to the α1β1 dimer (bottom, gray). (e) Comparison of the ribosome structure experimentally determined (middle), and two conformers (left and right) sampled by fluctuations along ANM mode 3 [62], approximating the ratchet-like rearrangement of the 50S subunit (green) with respect to the 30S (maroon), suggested by experimental data (see also the original work of [63]).
Molecular basis of pre-existing dynamics: insights from theory and simulations

The pre-existence of conformational motions is not unfamiliar to theoretical and computational biologists. Yet, their relevance to function is being established only recently.

Two major groups of studies have been undertaken toward this goal. The first uses analytical approaches (e.g., normal mode analysis (NMA) with coarse-grained (e.g., elastic network) [3,4]). NMA readily provides a hierarchy of accessible modes of motion uniquely defined by the given architecture. The low frequency modes among them are usually cooperative (collectively involving large segments/domains/subunits). They also are most readily accessible, since they are, by definition, along the direction of the lowest ascent (smallest curvature/force constant) away from the original energy minimum approximated by a harmonic well. As a result, they conceivably present a favorable, or intrinsically preferred, mechanism for dissipating the energy increase arising from external perturbations.

Two classical examples illustrating the relevance of low frequency modes to functional changes in structure are shown in Figure 1, panels (c) and (d). Panel (c) displays the open (left) and closed (middle) forms of E. coli adenylate kinase (AK), both determined by X-ray crystallography. The right diagram, on the other hand, displays an alternative conformation predicted by the anisotropic network model (ANM [17]) upon deforming the unliganded structure (left) along the slowest mode accessible in this substrate-free state [18]. Thus the closed form is readily approached by the open form upon movement along the direction of the first (energetically easiest) mode of motion. A similar phenomenon is observed in hemoglobin (Hb), a prime example of an allosteric protein (panel e). In this case, the tense (T) (left) structure is reconfigured to approximate the liganded (R2 (middle)) form, by moving along the second slowest mode direction (right) [19]. Similar results for other systems support the view of the intrinsic tendency of the unbound proteins to approach their bound conformations via low frequency modes [20,21,22]. An application to ribosome, illustrated in panel (e), also points to the functional relevance of slow modes. The low frequency modes predicted by coarse-grained NMA can now be readily retrieved and viewed for all PDB structures using recently developed web servers (e.g., [17]).

The second group of studies is based on MD simulations, accelerated by advanced algorithms such as replica exchange, steered MD or simplified models and force fields. While simulations require significantly longer computing time and may suffer from convergence problems, their main advantage is in exploring substrates that may not be accessible via NMA. Recent simulations clearly demonstrate the energetic grounds for the ability of AK to fluctuate between open and closed forms [23]. Likewise, unrestrained MD with implicit solvent shows the ability of HIV-1 protease to fluctuate between close, semi-open and fully open states, while predominantly populating the semi-open state [24]. Another extensive MD study of a series of protein–protein complexes showed the existence of an ensemble of conformers before binding, the stabilization of selected conformers—those distinguished by shape complementarity—to form complexes, succeeded by further structural rearrangements [25]. Simulations prove particularly useful when combined with NMA or coarse-grained models [26–31].

A key question, yet to be answered is how the internal dynamics of enzymes influences reaction rates. Theoretical studies on CypA dynamics indicate that a transfer of energy from first hydration shell and solvent-exposed regions into the protein interior all the way to the active site, through a network of vibrations, promotes catalysis by lowering the energy barrier along the reaction coordinate [32]. Similar observations were made for DHFR [1,32]. A link between collective mechanics and interactions at the catalytic site is suggested by the co-localization of the global hinge site with the catalytic site, shown for a representative set of enzymes using elastic network models [33]. Another recent ANM study showed the coupling between domain motions and loop rearrangements at the catalytic site of triosephosphate isomerase [34]. These studies also drew attention to the conservation of key residues that concert cooperative motions [1,32,33,35,36].

While the coupling between dynamics and function is recognized, quantitative assessment of the correlation between intrinsic dynamics and reaction rates remains to be clarified. The above described changes are physical events in the neighborhood of the original state. Catalytic reaction rates are controlled by chemical events farther along the reaction coordinate; their quantitative assessment requires consideration of free energy paths and barrier at transition state; and in this respect, the contribution of conformational motions to catalysis has been challenged [37]. The effect of distant mutations on DHFR catalytic reaction has been explained quantitatively in terms of the changes in the pre-organization of polar residues, arising as a consequence of the change in energy landscape in the presence of mutations [38].

Allosteric changes in conformations: Communication via concerted motions

Many allosteric biomolecules are multimeric and symmetric, which enhances the cooperativity of their conformational changes, leading to an all-or-none transition in accord with the MWC model [2**]. Yet, allostery appears to be featured by a broad range of proteins [39] including single domain proteins, and intermediates or sequential
events are also observed, especially in some small proteins and ligand-gated ion channels.

The prominent feature in all allosteric proteins appears to be the accessibility of functional conformers, and the efficient communication between distant residues (e.g., ligand-binding and catalytic residues). In nicotinic acetylcholine receptor (nAChR), for example, signal transduction occurs between nACh binding and pore opening sites separated by 50 Å. Such distant couplings may be readily accounted for by low frequency modes. The NMA of nAChR indeed showed that pore opening is ensured by the lowest frequency mode intrinsically favored by the structure irrespective of neurotransmitter binding [40]. This mode entails a global twist of the quaternary structure, similar to the pore opening mechanism identified for a series of potassium channels [41].

Low frequency modes are most effective in spreading signals, because of their cooperative nature and accessibility with minimal energy requirement. Yet, local effects may also play a critical role in mediating signals. The assessment of signal transduction mechanism may therefore necessitate the consideration of multiple modes, including those in the high frequency regime and mode–mode couplings [1,32,42,43]. For example, a loop (β4–α4) has been pointed out in CheY to mediate the rotational isomerization of substrate-binding tyrosine (Y106) succeeding the phosphorylation event 9.5 Å apart [26]. Subsequent simulations showed, however, that the loop is stabilized by the isomerization of Y106 in an active conformation [44], consistent with the population shift model, and MWC mechanism. Yet, a recently solved CheY shows the same loop in an intermediate conformation, inviting attention to the accessibility of intermediate conformations [45].

Signal transduction along well-defined pathways mediated by collective motions emerges now as a common feature of allosteric proteins [46,47]. Recent theoretical studies indeed demonstrate the occurrence of key sites with high allosteric potential [48,49] or strategically placed hot spot residues [30], and highlight pathways of allosteric signal propagation [50] consistent with experimental data [51]. These pathways predominantly involve conserved residues, as also deduced from sequence-based analysis [52,53], as illustrated in a recent MD study [54]. Recent work also shows how communication pathways relate to collective motions [55]. Notably, binding sites are reported to be located at regions that strongly affect the network of interactions, and such properties have been suggested to result from evolutionary pressure [56].

The spread of signals/perturbations via coupled motions, including in particular the most cooperative (lowest frequency) modes of motions, and conserved residues, thus

---

**Figure 2**

Combination of a pre-existing equilibrium and post-binding rearrangement. The protein (E) originally samples an ensemble of conformations, accessible via fluctuations in the global energy well (top panel), occasionally involving passages between substates within the well. The well is approximated by a harmonic potential (dashed curve) in coarse-grained NMA. Two conformations/substates are schematically shown, which are in a dynamic equilibrium before substrate (S) binding, as shown on the left box (bottom). The substrate (S) selects the conformer(s) that allows for optimal interaction (conformational selection). The stabilization of the final complex is subject to further rearrangement (induced fit). The energy profile/landscape of the protein E in the presence of the substrate differs from its unbound form (top right diagram) inducing a shift in the population of the substates in favor of the conformer that binds the substrate.
emerges as a plausible mechanism in allostery. This also suggests hinge sites in these modes to serve as ‘messengers’ for mediating allosteric communication, a conjecture to be tested by further studies.

Conclusion

Increasingly larger body of experimental and theoretical studies concurs on the ability of enzymes and/or allosteric proteins to sample, even in their inactive or substrate-free forms, conformations that approximate those required for biological function. This ability is structure-encoded.

Based on recent observations, can we reconcile the MWC and KNF models? The accumulating data appear to support the MWC model, in general. On the other hand, there exist examples where sequences of events or intermediate states are reported [45,47,57,58]. While the intrinsic dynamics pre-dispose the protein to bind its substrate, the final conformation is usually stabilized upon local rearrangements induced after binding [12**,16,21*,24*, 25*,26,39*,59], suggestive of the mechanism schematically described in Figure 2, and similar changes in energy landscape are entailed by allostery [60]. Therein a pre-existing equilibrium followed by the selection of the optimal conformation by the substrate and an induced fit to stabilize the final bound form is anticipated. Also, evidence has been presented for networks of coupled motions that facilitate enzyme catalysis, energy propagation or signal transmission [1*,30–32,35,36,42,50,51,55,56*], which may elicit a sequence of events, reminiscent of the KNF model. Clearly in some systems the sequential events may be too fast and subtle to be detectable by experiments resulting in an apparent all-or-none process dominated by a highly cooperative single mode of motion.

Irrespective of the kinetics of conformational change, the intrinsic ability of the protein structures to undergo conformational changes along directions that enable their function unequivocally supports the mapping structure → dynamics → function. Structures are usually accepted to be designable when they meet the criterion of being thermodynamically stable (the lowest energy conformer) for a diversity of sequences [61]. With the emerging functionality and robustness (insensitivity to structural and energetic details) of proteins intrinsic dynamics, the pre-disposition to perform functional changes in conformation may probably be viewed as another criterion for designing structures. Structures may have evolved to ‘move’ in the right directions.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest


Cyclophilin motions observed during catalysis are detected by NMR relaxation to be already present in the unbound (free) form of the enzyme. This suggests that the motions necessary for catalysis are an intrinsic property of the enzymes.


16. Fetler L, Kantrowitz ER, Vachette P: Direct observation of solution of a preexisting structural equilibrium for a mutant of
Substates populated by unbound receptor and ligand structures were reconciling the 'old' and 'new' 3-step mechanism for binding: diffusion, free conformer selection and selection of particular pairs to form complexes, some of which are tertiary between substates emerges as a major determinant of the


The potential of mean force for the reaction coordinate between the open and closed forms of AK is demonstrated to have a rather flat region between the open and relatively closed forms. These results imply that apo AK can fluctuate between the open and closed conformations in the absence of its substrate.


MD simulations of HIV-1 protease demonstrate that the protein exists in a dynamic equilibrium between its close (bound), semi-open and open (unbound) conformations in the absence of the ligand.


This study provides a quantitative measure based on KL divergence for the change in conformational dynamics induced by ligand binding.


An ensemble-based description of proteins is utilized to evaluate the extent to which perturbations at different sites propagate to the entire structure. Application to a database indicates that perturbations at binding sites affect the co-operativity between all residues. The authors suggest that binding residues are inherently located at sites poised to effectively propagate signals.


