Dear Author,

Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list.

For correction or revision of any artwork, please consult [http://www.elsevier.com/artworkinstructions](http://www.elsevier.com/artworkinstructions).

Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof. Click on the 'Q' link to go to the location in the proof.

<table>
<thead>
<tr>
<th>Location in article</th>
<th>Query / Remark: click on the Q link to go</th>
</tr>
</thead>
</table>
| Q1                  | Please complete and update the reference given here (preferably with a DOI if the publication data are not known): [64].  
                      For references to articles that are to be included in the same (special) issue, please add the words 'this issue' wherever this occurs in the list and, if appropriate, in the text. |
| Q2                  | Please verify the slight change made in the sentence: 'However, at lower ... recognition mechanism'. Please update Ref. [64]. |

Thank you for your assistance.
Allostery and population shift in drug discovery

Gozde Kar¹, Ozlem Keskin¹, Attila Gursoy¹ and Ruth Nussinov²,³

Proteins can exist in a large number of conformations around their native states that can be characterized by an energy landscape. The landscape illustrates individual valleys, which are the conformational substates. From the functional standpoint, there are two key points: first, all functionally relevant substates pre-exist; and second, the landscape is dynamic and the relative populations of the substates will change following allosteric events. Allosteric events perturb the structure, and the energetic strain propagates and shifts the population. This can lead to changes in the shapes and properties of target binding sites. Here we present an overview of dynamic conformational ensembles focusing on allosteric events in signaling. We propose that combining equilibrium fluctuation concepts with genomic screens could help drug discovery.

Addresses

¹Center for Computational Biology and Bioinformatics and College of Engineering, Koc University Rumelifeneri Yolu, 34450 Sariyer Istanbul, Turkey
²Basic Science Program, SAIC-Frederick, Inc., Center for Cancer Research Nanobiology Program, NCI-Frederick, Frederick, MD 21702, United States
³Sackler Inst. of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

Corresponding author: Nussinov, Ruth (ruthnu@helix.nih.gov)

Current Opinion in Pharmacology 2010, 10:1–8

This review comes from a themed issue on Special new technologies the importance of protein dynamics
Edited by Dennis R Livesay

1471-4892/$ – see front matter
Published by Elsevier Ltd.


Introduction

A major aim in drug discovery is to identify protein targets whose inhibition can result in disease treatment. Discovery can be top-down or bottom-up: Top-down often involves a physiology-based approach; it focuses on disease phenotypes and (initially) foregoes direct target identification and mechanistic understanding. By contrast, a bottom-up strategy follows a target-based paradigm; as such, it benefits from genomic data. Functional genomics employs large-scale exploration to figure out regulatory networks, cellular pathways, forward-regulation and backward-regulation, and signal transduction. It aims to understand disease mechanisms. Nonetheless, a key question in drug discovery is how the abundance of proteomic data can be used toward effective therapeutic strategy under given conditions. Here we argue that to be useful, proteomic data and genomic screens should be combined with current concepts in structural biology that relate to the fundamental role of dynamic conformational ensembles in molecular recognition. Together, these could lead to a mechanistic understanding of protein function on the molecular level and provide comprehensive strategies in drug discovery. Key to successful discovery is accounting for dynamic changes in the cellular environment; such changes and cellular response are largely reflected in signaling. Signal transduction relays extracellular information to the interior of cells. Cell surface-membrane-bound receptors are stimulated by ions, small molecule morphogens, hormones, neurotransmitters, and covalent modification events. This initiates a cascade of intracellular protein interactions that transmits to genome activation and specifies cellular expression. Proteins are dynamic and this is particularly the case for signaling proteins, which are frequently disordered [1]. Here, we describe proteins in terms of their dynamically fluctuating conformational ensembles and ensemble distributions; how these distributions change upon structural perturbations, such as those caused by binding of proteins, cofactors, DNA, small molecules or drugs [2]; or changes in the cellular environment; and on the ramifications of these fundamental phenomena, called allosteric effects [3,4,5,6,7,8,9,10], toward drug discovery [11,12,13,14,15].

Conformational ensembles and population shift

In solution, proteins exist in an ensemble of conformational substates around their native states that are in equilibrium [16]. The substates are separated by low barriers, and their populations follow a statistical thermodynamic distribution [17–20]. The heights of the barriers reflect the timescales of the conformational exchanges.

The conformer whose binding pocket shape is most complementary to the ligand conformation will be selected for binding [21]. The equilibrium will then shift toward that bound protein conformation. The conformer selected may not have the lowest energy; however, binding will stabilize it leading to ‘population shift’ toward this conformer. Such a description [19,20] is fundamental to the understanding of binding mechanisms: it argues that if we accept that proteins pre-exist in a broad range of conformational states, among these there will be states with binding site shapes that are complementary to those of many possible ligands. This ‘conformational selection and population shift’ model proposed over a decade ago
2 Special new technologies the importance of protein dynamics

[18–20] as a primary binding mechanism contrasts the ‘induced fit’ hypothesis [22]. According to the induced fit model, static unbound proteins accommodate an incoming substrate by flexibly adapting their ligand binding site shape from an ‘open’ to a ‘closed’ state, with the adaptation induced by the ligand. On the contrary, the conformational selection and population shift model posits that the closed state already exists in solution in the free (unbound) protein state; since it is complementary to the ligand it selectively binds, with subsequent induced fit on a minor, local scale to optimize side-chain interactions. The kinetic differences between conformational selection and population shift versus induced fit scenarios [23**] reflect ligand concentrations. If extremely high—induced fit may prevail. However, at lower physiological level concentrations, conformational selection with population shift is expected to be the major molecular recognition mechanism. Induced fit takes place on much faster timescales than conformational selection. This is because the complementary conformer has lower population; i.e. it is a higher energy state. Consequently, binding timescales via conformational selection reflect the barriers-crossing. On the contrary, in induced fit, in principle any conformer can bind; hence the faster rate. Since the population of the complementary conformer is low, direct experimental data on conformational selection has been difficult to obtain [24**]. This has led to the mistaken view that since experiment only sees the open state and the ligand-bound state, recognition takes place via induced fit. However, currently NMR is increasingly able to detect such low population states and provide direct validation of the conformational selection and population shift model [24**,25**,26**,27**,28,29]. From the theoretical standpoint, this view is based on the recognition that the free energy landscape of proteins is dynamic [18]. Conformational substates populate shallow wells with low barriers. Binding events change the relative stabilities of the substates; thus altering the landscape [30]. Equilibrium shift toward the bound conformer then takes place, continuing the binding. Kern and her colleagues showed in exquisite detail how population shift can take place [27**] via transient interactions in high energy states; such states were also visualized in the amino-terminal processing of the HIV-1 protease [31]. Transient interactions that result from internal protein motions help the protein get over energy barriers. Figure 1 illustrates the population shift concept.

**Allostery reflects population shift**

Population shift of dynamic conformational ensembles is the origin of the allosteric effect: the perturbation that takes place during the binding of an effector at a site other than the active site (the so-called allosteric site), propagates in the structure, leading to conformational changes at the active site. The changes can be minor; nonetheless, they can affect ligand selectivity and affinity. Allosteric perturbation can arise from binding of small or large molecules; from changes in temperature, pH, concentration, or ionic strength; and from covalent changes such as post-translational modification [32,33] or mutational events [33]. The energetic strain at the perturbation site dissipates in the structure like waves, and transmits to the active site. Propagation is via multiple pathways through dynamic changes in atomic contacts, where pathways can be viewed as summations of individual dynamic micro-pathways [5**]. Since perturbation can involve enthalpic and entropic components, the outcome may be reflected in changes of shape (enthalpy); changes of shape and atomic fluctuations (enthalpy and entropy) or only entropy changes [8**]. If the shape changes are minor, the process is described as entropy-dominated; if more pronounced it is enthalpy-dominated (Figure 2). Binding can be a function of concentration or selectivity; selectivity reflects shifts of the ensemble of substrates by prior allosteric perturbations. Allostery is a cooperative event, upregulating or downregulating protein activity. From the functional standpoint, the key role of allosteric events is to increase binding selectivity at the target site [34]: binding to even slightly different allosteric effectors or at different allosteric sites can enhance specificity. From the pharmacological standpoint, allosteric effects can adversely affect protein function: disease-related
Allostery and population shift in drug discovery Kar et al.

Illustration of the conformational changes in allosteric proteins. Known inactive and active structures for four signaling proteins (CheY, Rap2a, Cdc42, and IRK) are obtained from the PDB. Inactive (in pink color) structures are superimposed onto active (cyan) structures. The superposition is based on matched residues with the distance between superimposed Cα atoms <2 Å. The conformational changes (unmatched residues) are highlighted in red and blue, respectively, for the inactive and active structures. The classification of the conformational changes is based on Tsai et al. [8†]. (a) CheY (PDB IDs: 3chyA and 1fqwA) is classified as showing no conformational change; (b) Rap2a (PDB IDs: 1ka0A and 2rapA) subtle conformational change; (c) Cdc42 (PDB IDs: 1an0A and 1nf3A) minor conformational change; (d) IRK (PDB IDs: 1irkA and 1ir3A) large conformational change.

mutations often lie on major allosteric routes [5**,35**]. The efficacy of drugs that bind residues on major propagation pathways can be expected to be higher. Hence, identification of major pathways in the ensemble is an important goal.

Conformational ensembles: implications in pharmacology

Shared binding sites in signaling proteins have many pre-existing shapes. Since prior binding or covalent change events re-distribute the ensemble, the population of certain binding site shapes will increase; which shape predominates reflects the prior event and the perturbation it elicits. The partner is an ‘allosteric effector’ and the consequent population shift is ‘allostery’. Key to the understanding of allosteric shift and its ramifications for drug discovery is that such shift and thus active site shape may enhance substrate binding or disfavor it. Disease-related mutations are similarly allosteric effectors shifting the equilibrium toward a binding-disfavored state. Compensatory ‘rescuing’ mutations re-distribute the population toward the ligand binding-favored state. The goal of allosteric drugs [12**,15] is to similarly alter the distributions of the states, to shift them to the native functional state. The conformational changes elicited by the drug at the binding site may be minor; nonetheless, they may allosterically lead to subtle changes in active site side-chains. Allosteric drugs do not merely inhibit; similar to other allosteric events, they can enhance, and even activate enzymes [36**]. From the conformational standpoint, the mechanism is unchanged: propagation of a perturbation caused by (drug) binding far away.

Combining free energy landscapes with genome-level information can help drug discovery: case studies

Diseases are complex, and identifying drugable, disease-relevant proteins and accounting for their role in the network is a challenging task. Genomic screens [37] can identify inherited disorders, and genome-wide expression profiling can provide the differences in dynamic gene expression patterns between normal and diseased states as for example in leukemia [38] and transient populations that emerge in specific stages of development [39]; as such they assist in identifying protein drug targets. Yet, while essential, structural determination of these proteins is insufficient; their interaction partners, where and how they interact and the allosteric conformational changes that these binding events elicit are also needed. These describe proteins in terms of their role in the network; and at the same time they can allow atomic-level drug design. The first step involves modeling the structural proteome; the second modeling the conformational changes. Modeling of the structural proteome can provide the complete static structural network; on the contrary, the free energy landscape complements this description by helping in understanding the dynamic changes in the distributions of the substates (Figure 1) many of which are functionally relevant. These are important because they can explain key biological observations such as switching mechanisms that result from allosteric events. Below, we present examples of allosterically induced conformational changes resulting from population shift following binding events. The examples focus on signaling proteins. Signaling pathways are complex and dynamic and are important for identifying possible therapeutic targets. The examples are taken from the protein databank (PDB) and from structural prediction of protein-protein interactions using Prism 4.0L. Since the PDB contains a limited number of experimental complexes, a reliable modeling tool, which allows large-scale application, is essential. Prism employs a highly efficient strategy to predict protein associations.
on the proteome scale, to construct pathways [41] and characterize networks [42]. The Prism rationale argues that if any two proteins contain regions on their surfaces that are similar to complementary partners of a known interface, in principle these two proteins can interact with each other through these regions. Since the number of distinct binding motifs is limited in nature [43] and structurally different proteins can interact via similar interface architectures [44], such a knowledge-based strategy, which utilizes structural and evolutionary similarity, is powerful. This approach is made more physical and biologically relevant by including flexibility and energetic assessment in the modeling using FiberDock [45], a flexible docking refinement server.

Allosteric Ras proteins in signaling pathways

Ras proteins mediate signaling pathways that control cell growth and differentiation. They act as molecular switch by cycling between active guanosine triphosphate (GTP)-bound and inactive guanosine diphosphate (GDP)-bound states. They are activated by nucleotide exchange factors Son of sevenless (Sos) and Ras guanine nucleotide releasing factor 1 (RasGRF1) upon conversion of GDP-bound Ras to GTP-bound. Their large conformational changes are best described by the population shift model rather than induced fit [46]. In addition to the active and inactive states highly populated intermediates are also sampled [42]. In its active state, Ras can interact with effectors in signaling cascades, whereas inactive Ras cannot [48]. In the mitogen-activated protein kinase (MAPK) signaling pathway, Ras is first activated by Sos, then it binds Raf protein kinases (A-Raf, B-Raf, and C-Raf). Although there are binding studies on Ras interaction with C-Raf [49,50], little is known on binding of B-Raf to Ras. Since B-Raf is important in many cancer types, details of its interactions are essential for drug design. Here, using the active and inactive conformations, we search for possible interactions of Ras with B-Raf using Prism [40]. Results indicate that active Ras can favorably bind to B-Raf while inactive Ras cannot owing to steric effects. Figure 3a illustrates the conformational change upon activation. Figure 3b displays the Prism-predicted Ras-B-Raf interaction. When binding to B-Raf, Ras uses switch I (residues 30–38) and switch II (residues 60–76) regions that correspond to the most significant conformational change. The figure indicates that without this conformational change, inactive Ras-B-Raf interaction is unfavorable.

Allosteric inhibition and allosteric drugs

Misregulation of signaling pathways often ends in disease. Drug design generally focuses on active-site inhibitors. However, active-site topologies are usually conserved in families [51], thus drugs targeted to active sites can lead to side effects. Allosteric sites can be specific [52] since their conformational details may be less conserved. An allosteric antibody (Fab40) was designed to inhibit serine protease hepatocyte growth factor activator (HGFA) [52**]. HGFA promotes cancer by activating pro-hepatocyte growth factor and triggering the HGF/Met signaling pathway [54]. The HGFA-Fab40 complex revealed that Fab40 binding was not accompanied by major conformational changes other than of HGFA 99-loop [53**]. The 99-loop movement is an allosteric switch regulating HGFA activity: upon Fab40 binding the equilibrium is shifted away from the functionally active state [53**]. The allosteric inhibition of HGFA is illustrated in Figure 4a. A mutant of Fab40 restores the activity.

Rather than focusing on a single conformational state, allosteric inhibitors could advantageously target multiple states [55**]. Although progress has been made [56], conformational changes in the ensemble are not equally druggable [57] owing to differences in energetics and binding pockets’ accessibility [24**]. A well known example of an allosteric inhibitor that targets the inactive conformation is imatinib (Gleevec) that inhibits Abl tyrosine kinase associated with chronic myelogenous leukemia. Imatinib binding to inactive Abl tyrosine kinase dynamically shifts the free-energy landscape to favor the imatinib-bound inactive conformational substrate. A schematic representation of the free energy landscape and population shift for allosteric inhibition of Abl tyrosine kinase is visualized in Figure 4b.
Fluctuations are important for function

Finally, it behooves us to note that equilibrium fluctuations take place over a broad range of timescales, and relate to function at different levels [58]. Examples include the fluctuations of the quaternary structure of the small heat shock protein with its clients that are important for its role in the cellular protein homeostasis network [59] and protein degradation by the 20S core-particle proteasome via gated substrate access by interconversion between conformations that extend inside (closed gate) or outside (open gate) on timescales of seconds [60]. Conformational dynamics can provide clues to functionally relevant motions that determine the enzyme turnover rate [61]. Adenylate kinase provides a striking example [28,62]. Small amplitude nanosecond motions (under 10 ns) do not complete open to closed state switches that are rare events on longer timescales [61]. These can take place via local unfolding that results from global motions [63]. Allosteric communication may explain the LID and NMP domain closure and substrate turnover [64]. In dihydrofolate reductase millisecond timescale fluctuations are exquisitely tuned for every intermediate in the catalytic cycle [25]. Free enzyme dynamics already encompass all the confor-
6 Special new technologies the importance of protein dynamics

mations that are necessary for substrate binding, preorganization, transition state stabilization, and product release, and conformational selection and substrate population shift at each catalytic step can accommodate specificity and efficiency [44].

Conclusions
Proteins exist in conformational ensembles around their native states, many of which are functionally relevant [18–20,65]. The ensemble can present a large number of substates. The population of each substate is not static; it changes dynamically with the conditions. This dynamic landscape is the outcome of environmental fluctuations that physically perturb the protein structure. The strain energy dissipates in the structure by radiating out via major and minor pathways. Perturbation at the allosteric site can be in the form of a conformational change; or if there is no conformational change, the stiffness (entropy loss) at the binding site can propagate dynamically. This is how allosteric drugs work. In vivo, at any moment a combination of allosteric events takes place. For example, along with drug binding external cellular conditions change; disease-related mutations occur; cofactors bind; post-translational modifications, such as phosphorylation, or acetylation take place [66]. This is particularly the case for the multi-modular signaling proteins that have to respond to changes in subcellular distribution, enzymatic activity and multiple binding events. These present co-occurring allosteric events that would shift the populations, altering the relative distributions of the substates. Pathways that initiate at several perturbation sites merge similar to waves, to enhance or dissipate deformations at the active site.

From the standpoint of allosteric drug discovery, considering the complexity of the forward and backward regulation in signaling networks [67] and the cooperativity of functional interactions, it is difficult to predict the biological consequences of therapeutic intervention. Not only is it difficult to model the detailed regulatory sequence of events; the multiple binding events are expressed in shifts in the distributions of conformational substates presenting complex free energy landscapes, making the prediction of the conformational changes in a target binding site an extremely challenging task. At the same time, allosteric drugs allow modulation of signals and responses, in contrast to drugs binding at active sites; as such, they hold great promise for future developments.

Acknowledgements
This work has been supported by TUBITAK (Research Grant Numbers: 109T343 and 109E207). This project has been funded in whole or partly with Federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This research was supported (partly) by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

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


This review article argues that there are multiple, pre-existing allosteric pathways and disease-related mutations usually lie on major propagation routes. The authors provide several experimental examples substantiating this argument.


This paper emphasizes that allosteri may not necessarily involve a change of shape and provides a definition for allosteri in thermodynamic terms: allosteri can be mediated by enthalpy, by enthalpy and entropy, or solely by entropy.


The authors perform molecular dynamic simulations to investigate binding of a cancer drug, imatinib to several tyrosine kinases. Binding free energy calculations indicate that conformational selection governs imatinib binding to kinases Abl, c-kit, and Lck.


In this paper the authors present affinity and efficacy in terms of vectorial energy flow between a modulator, conduit (the receptor protein), and guests that are other molecules such as receptors, membrane-bound proteins, or signaling proteins in the cytosol. Vectorial energy flow is the energy propagating from the perturbation site. See also Leitner (Ref. [7]).


This paper considers allosteric effects in terms of the affinity and the efficacy. It describes agonist response as an allosteric reduction of affinity and/or efficacy. It argues that the antagonist will have different properties depending on which of these is primarily affected.

Allostery and population shift in drug discovery Kar et al. 7


This paper compares induced-fit and conformational selection mechanisms from the perspective of kinetics; the kinetic rate constants can determine which mechanism will dominate. If ligand concentration is extremely high, induced-fit may dominate; however, at lower physiologically relevant ligand concentrations, conformational selection will be the preferred kinetic pathway.


This paper provides many examples that illustrate the role of conformational selection and population shift in a broad range of macromolecular binding events, including protein-ligand, protein-protein, protein-DNA, protein-RNA, and RNA-ligand interactions.


The authors used NMR relaxation dispersion to investigate the dynamics in the dihydrofolate reductase enzyme. Their results suggest that dynamics in this enzyme are exquisitely tuned for every intermediate in the catalytic cycle highlighting the optimization of structural fluctuations for efficient catalysis and regulation.


The authors elegantly unravel interconverting substates of the human proline isomerase, cyclophilin A based on crystallographic approaches in combination with NMR analysis.


Using nuclear magnetic resonance (NMR) dynamics and molecular simulations, the authors inferred a molecular pathway for the conformational transition in a signaling protein, nitrogen regulatory protein C (NtrC). The molecular pathway involves transient, non-native hydrogen bonds that play a key role in stabilizing the transition state. However, additional transient interactions may also take place.


8 Special new technologies the importance of protein dynamics


56. This paper discusses the advantages of trapping macromolecules in their inactive state and provides examples of allosteric drugs and their mechanisms. The authors highlight the use of nuclear magnetic resonance spectroscopy in drug discovery and design studies.


61. The authors applied a quantitative mass spectrometry method to decipher the complex array of the small heat shock protein (sHSP) dynamic assembly/disassembly and client-binding events. They find that thermally regulated quaternary dynamics of the sHSP establish and maintain the plasticity of the system. The findings indicate the importance of the equilibrium fluctuations of the sHSP's quaternary structure with its clients for its role in the cellular protein homeostasis network.


63. Using methyl-transverse relaxation optimized nuclear magnetic resonance spectroscopy, the authors show that gating termini of the 20S core-particle proteasome are in conformations that extend either inside (closed gate) or outside (open gate). Dynamic regulation of 20S core-particle proteasome activity takes place by interconversion between these conformations.


65. The authors question the key issue whether the protein moves by specific hinges or it cracks and becomes locally unfolded during its structural change. They provide a comparison of the studies investigating conformational dynamics around the equilibrium basin of the enzyme adenylate kinase. See also Whitford et al. Ref. [63].


70. This paper reviews recent experimental and computational techniques that describe native state dynamics. The authors also provide some cases where native dynamics can be detected experimentally and interpreted in the light of functional landscapes paradigm.


75. The authors present a remarkable model for DFG flipping in Abl kinase based on molecular dynamic simulations, crystallographic, and kinetic experiments. The model suggests a pH-dependent DFG flip modulated by electrostatic changes inherent to the catalytic cycle that is supported experimentally by Abl-imatinib binding assays.