Disordered allostery: lessons from glucocorticoid receptor

Hesam N. Motlagh¹ · Jeremy A. Anderson¹ · Jing Li¹ · Vincent J. Hilser^{1,2}

Received: 22 October 2014 / Accepted: 5 April 2015 / Published online: 23 April 2015

© International Union for Pure and Applied Biophysics (IUPAB) and Springer-Verlag Berlin Heidelberg 2015

Abstract Allostery is a biological regulation mechanism of significant importance in cell signaling, metabolism, and disease. Although the ensemble basis of allostery has been known for years, only recently has emphasis shifted from interpreting allosteric mechanism in terms of discrete structural pathways to ones that focus on the statistical nature of the signal propagation process, providing a vehicle to unify allostery in structured, dynamic, and disordered systems. In particular, intrinsically disordered (ID) proteins (IDPs), which lack a unique, stable structure, have been directly demonstrated to exhibit allostery in numerous systems, a reality that challenges traditional structure-based models that focus on allosteric pathways. In this chapter, we will discuss the historical context of allostery and focus on studies from human glucocorticoid receptor (GR), a member of the steroid hormone receptor (SHR) family. The numerous translational isoforms of the disordered N-terminal domain of GR consist of coupled thermodynamic domains that contribute to the delicate balance of states in the ensemble and hence in vivo activity. The data are quantitatively interpreted using the ensemble allosteric model (EAM) that considers only the intrinsic and measurable energetics of allosteric systems. It is demonstrated that the EAM provides mechanistic insight into the distribution of states in solution and provides an interpretation

for how certain translational isoforms of GR display enhanced and repressed transcriptional activities. The ensemble nature of allostery illuminated from these studies lends credence to the EAM and provides ground rules for allostery in all systems.

Keywords Allostery · Intrinsic disorder · Glucocorticoid receptor · Protein stability · Ensemble

Introduction

Allostery is a near ubiquitous biological phenomenon that is vital to many biological processes, endowing cells with the ability to alter their function in response to subtle chemical signals (Nussinov et al. 2013). Structural interpretation of allosteric mechanism was originally proposed to describe hemoglobin (Perutz 1970; Perutz et al. 1998), and this paradigm has been shown to be applicable to many systems that are amenable to structural characterization (Daily and Gray 2009). In the past decade, allostery has undergone a resurgence, with studies focusing on the importance of dynamic and intrinsically disordered (ID) proteins (IDPs) or regions (IDRs) in facilitating allostery (Cui and Karplus 2008; Motlagh et al. 2014). Recent experimental evidence has mounted to support the unstructural view of allostery including coupled folding and binding (Ferreon et al. 2013), low- and high-affinity conformational switching of binding sites (Garcia-Pino et al. 2010), and modulation of binding affinity by changes in dynamic fluctuations in proteins (Popovych et al. 2006). What has become clear from these studies is that allostery can no longer be explained exclusively in structural terms alone. Instead, allostery should be treated as a dynamic continuum of potential mechanism that nature can use to propagate the allosteric signal (Fig. 1).

This article is part of a Special Issue on “The Role of Protein Dynamics in Allosteric Effects” edited by Gordon Roberts.

✉ Vincent J. Hilser
Hilser@jhu.edu

¹ T.C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD 21218, USA

² Department of Biology, Johns Hopkins University, Baltimore, MD 21218, USA

Fig. 1 The dynamic continuum of allosteric phenomena. Allosteric mechanisms that exhibit increase in dynamics or disorder (i.e., increasing towards the *bottom*). Examples include rigid body motions (human hemoglobin (Perutz 1970; Perutz et al. 1998)), sidechain and backbone dynamics (catabolite activator protein (Popovych et al. 2006)), local unfolding (aminoglycoside N-(6′)-acetyltransferase-li [AAC] (Freiburger et al. 2011)), and intrinsic disorder (glucocorticoid receptor (Li et al. 2012) and alpha-synuclein (Sevcsik et al. 2011)). Figure adopted from (Motlagh et al. 2014)

The question has remained whether it is possible to quantitatively describe allostery in general terms that are transferable between the different regions of the dynamic continuum (Fig. 1). A recently proposed ensemble allosteric model (EAM) attempts to answer this question by describing allostery in terms of the intrinsic coupled equilibria of proteins (Hilser and Thompson 2007; Hilser et al. 2012; Motlagh et al. 2014). This model is capable of describing structured, dynamic, and even disordered systems. This is of particular importance in the field of IDPs, whose members, by definition, lack a well-defined stable tertiary structure in their native state. The biological significance of IDPs/IDRs can be appreciated by their prevalence in eukaryotes, more specifically their hyper-abundance in transcription factors (Liu et al. 2006). IDRs confer versatility required by cells that use a limited set of players to integrate multiple signals giving them the ability to interact with different binding partners with high specificity and low-affinity (Smock and Gierasch 2009). It has even been demonstrated that certain IDPs can switch binding cooperativity dependent on what other protein partners are currently bound (Ferreon et al. 2013). What has become clear is that nature utilizes the full repertoire of dynamics and disorder to relay allosteric information.

In this chapter, we discuss general models of allostery and the emerging role of conformational entropy in driving allosteric signals. We then focus on the application of the EAM to explain data from the intrinsically disordered domain of the human glucocorticoid receptor (GR), a member of the steroid hormone receptor (SHR) family. This disordered domain has multiple translational isoforms that fold cooperatively to a state or set of states that are thermodynamically similar to a globular protein. It has been demonstrated that these isoforms and disordered regions correspond to thermodynamically coupled domains and that this coupling is a contributing factor to *in vivo* activity. Lastly, it is then demonstrated that these results can be quantitatively reconciled using a two-domain EAM. These studies have broad implications in the SHR family specifically, and in the IDP/IDR field in general.

Models of allostery and the emerging role of entropy and disorder

Allostery is fundamentally rooted in that binding of an effector molecule is physically removed from the site it affects (Cui and Karplus 2008). This action at a distance is why a general, quantifiable, and transferable allosteric model has remained elusive. The two classic models of allostery are the MWC (Monod, Wyman, and Changeux (Monod et al. 1965)) and the KNF (Koshland, Nemethy, and Filmer (Koshland et al. 1966)), which emerged around the same time when the first high-resolution structures of proteins were being elucidated (Muirhead and Perutz 1963; Perutz et al. 1960). These timely models proved extremely useful in rationalizing and quantifying the allostery of many systems with the MWC model emerging as the dominant model (Changeux 2005).

The MWC model posits that the macromolecule is in equilibrium between a low-affinity tensed (T) state and high-affinity relaxed (R) state. The macromolecule can bind ligand at multiple sites, and the allosteric process is driven by the T- and R-states having different binding affinities. A fundamental limitation of MWC is that it does not explicitly consider interactions between subunits and assumes the conformational change to be an “all or nothing” positively cooperative process. The KNF or Adair (1926) model explicitly considers interactions between subunits and allows for intermediate conformations where adjacent subunits sense one another. Due to the interaction parameter in the KNF model, it can be applied to both positively and negatively cooperative allosteric binding processes depending on the value of the interaction parameter. Despite their generality and applicability to multiple systems, both models are limited by not addressing “how” binding drives allostery. In particular, the recent emergence of allostery driven by dynamics and intrinsic disorder is difficult to reconcile within the context of structure-based allosteric pathways and the MWC/KNF models.

The quantitative role of dynamic allostery was first explored in 1984 by Cooper and Dryden (1984). By considering the intrinsic energetics of a molecule subject to thermal fluctuations, it was found that frequencies and amplitudes can change around the mean atomic positions resulting in allostery on the order of a few kcal/mol without changing the ground state structure—i.e., without changing the crystal structure of a protein, effector binding can have an allosteric response purely from conformational entropy. This result proved prescient as multiple systems have been demonstrated to manifest this process. Direct measurements of conformational entropy proved elusive until recent advancements in biomolecular NMR that led to an empirical “entropy meter” (Wand 2013). These measurements have been recently applied to the catabolite activator protein (CAP) and calmodulin demonstrating the role of conformational entropy in biological processes (Popovych et al. 2006; Tzeng and Kalodimos 2009, 2012). Specifically, in CAP it was found that, without a change in the ground state structure, significant change in conformational entropy of the side chains upon ligand binding drives the allosteric process (Kalodimos 2012). Reconciling these results in the context of a quantitative model remains notional. An important area for future research is to decipher the contribution of conformational entropy to allostery in a quantitative manner and to apply these models to systems where measurements can be made.

More recently, the realization of allostery in disordered systems has become apparent (Ferreon et al. 2013; Garcia-Pino et al. 2010). Some systems have even demonstrated that allosteric processes can switch from positive to negative cooperativity depending on which ligands are present (Ferreon et al. 2013); a result which is difficult if not impossible to reconcile structurally. The culmination of data for dynamic and disordered driven allostery requires consideration of the statistical nature of the interaction between coupled regions of the molecule. In the following sections, we focus on the application of the EAM to quantitatively describe coupling and allostery in disordered domains of GR.

Allosteric coupling in the intrinsically disordered domain of glucocorticoid receptor

Allostery is of particular importance in transcription factors that must respond to and integrate multiple signals in the cell. The prevalence of IDPs and IDRs in transcription factors suggests that disorder can somehow facilitate allosteric coupling. Unfortunately, experimental data supporting this hypothesis has been sparse until recently (Ferreon et al. 2013; Garcia-Pino et al. 2010; Li et al. 2012). Given that dysregulation of transcription factors can lead to disease (Nussinov and Tsai 2013), it is of critical importance to study how intrinsic

disorder contributes to the allosteric signal and hence biological function.

The SHR family of transcription factors is an ideal set of systems to study the interplay between allostery, intrinsic disorder, and function. SHRs consist of a conserved domain architecture consisting of an ID N-terminal domain (NTD), a DNA binding domain (DBD), and a C-terminal ligand binding domain (LBD) (Fig. 2a) (Lavery and McEwan 2005). Strikingly, there is no sequence conservation between the NTDs of the SHR family yet they all contain an activation function (AF) region required for full transcriptional activity (Lavery and McEwan 2005). In GR, the first AF region (AF-1) is found in the disordered NTD. For the human GR, eight different translational isoforms have been reported that vary in length, tissue distribution, and in vivo activity (Fig. 2b) (Lu and Cidlowski 2005). Interestingly, the start sites of some isoforms of the NTD of GR are outside of the AF-1 region suggesting biological pressure to maintain these sites (Fig. 2b). A perplexing question has been what functional roles do the conserved start sites outside of the AF-1 region serve? Previous work has revealed that coupling exists between different ID regions outside of the AF-1; specifically, the N-terminal and C-terminal regions flanking the AF-1 are negatively coupled as determined by in vivo activity (Kumar and Thompson 2010). Interestingly enough, it was even shown that the ID-NTD is also coupled to the structured DBD such that structure is gained in the AF-1 that preferentially binds certain co-regulatory binding partners (Kumar et al. 1999). These results qualitatively indicate that there is coupling and suggest that gaining structure in the ID-NTD is required for biological function, yet a quantitative measurement of these couplings and a predictive model for folding remain elusive.

It is well-established that some IDPs undergo coupled folding and binding and that the folded state is the functional conformation (Smock and Gierasch 2009; Tantos et al. 2012). Indeed, cooperative folding of the AF-1 of GR with the osmolyte trimethylamine *N*-oxide (TMAO) increases its affinity for known binding partners as well as protection from protease digestion (Kumar et al. 2001). The A, C2, and C3 isoforms of the GR-NTD also fold cooperatively in the presence of TMAO, a hallmark of a spontaneously folding protein (Fig. 3a) (Li et al. 2012). What is clear from analysis of Fig. 3a is that the three isoforms in question have different stabilities and that removal of the residues between A and C3 increases the stability of the C3 isoform (Fig. 3b). Moreover, when the in vivo activities of these isoforms are plotted against the measured stabilities, there is a clear relationship (Fig. 3b). Taken together, these results support three hypotheses: (1) that the GR-NTD undergoes coupled folding and binding to perform its biological role, (2) that a contributing factor to this process is the energetic cost of folding, and (3) the different isoforms can modulate the stability and hence biological

Fig. 2 Glucocorticoid receptor domain architecture and naturally occurring translational isoforms. **a.** Depicted from N-terminus to C-terminus is the domain architecture of GR. From left to right is the intrinsically disordered N-terminal domain (NTD), the DNA binding domain (DBD), and the ligand binding domain (LBD). The NTD consists of two coupled domains: the regulatory (R) region and the functional (F) region in red and yellow, respectively. **b.** The eight naturally occurring

translational isoforms of GR are depicted in decreasing length from *top* to *bottom* with *numbers in parentheses* indicating the methionine amino acid number that starts from isoform. Nomenclature is as defined before (Lu and Cidlowski 2005). The activation function 1 (AF-1) domain is indicated on *top* of the isoforms as amino acids 77–262 (Lu and Cidlowski 2005)

activity of this process (Li et al. 2012). This led us to propose the model that the NTD consists of at least two thermodynamically coupled domains: a regulatory (R) region and a functional (F) region (Fig. 2a) (Li et al. 2012).

Closer analysis of the TMAO induced refolding data reveals that there is a relationship between the *m*-value and the length of the protein (Fig. 3c). The *m*-value describes the stability of the protein and its dependence on the concentration of denaturant or osmolyte (Street et al. 2008):

$$\Delta G_{\text{protein}}(x) = \Delta G_{\text{H}_2\text{O}} - m^* [x] \quad (1)$$

where $\Delta G_{\text{protein}}$ is the stability at a given concentration of osmolyte/denaturant, $\Delta G_{\text{H}_2\text{O}}$ is the stability of the protein in the absence of osmolyte/denaturant, $[x]$ is the concentration of osmolyte/denaturant, and *m* is the *m*-value as described. Interestingly, the GR-NTD isoforms fall on the same line as globular proteins. Given that the *m*-value from TMAO corresponds primarily to backbone burial (Auton and Bolen 2005), this result suggests that the refolded conformation of the GR-NTD isoforms is thermodynamically similar to a globular protein fold. In other words, when GR folds, it buries as much backbone surface (on a per residue basis) as globular proteins. The fact that the full length and different length truncated isoforms fall on the same line suggests that not only the F region (C3-NTD) folds to a globular structure but that the R

region (included in the A-NTD) does as well (Fig. 3c, gray circles).

Although the model proposed in Fig. 2a reconciles the data, can it be used to describe GR function quantitatively? What is clear is that the R and F domains are coupled and that each domain can access at least a folded and unfolded state (Fig. 3). Given this information, we can formulate a model that takes into account the intrinsic energetics and also manifests the conformational heterogeneity of the system. In the following section, we will develop a formalism using the EAM and then fit the data accordingly.

Development of a two-domain ensemble allosteric model

It is clear from the dynamic continuum of allostery (Fig. 1) that a quantitative formalism to describe all the possible allosteric phenomena will require an approach that can account for the observed conformational heterogeneity found in disordered systems. Although the limiting case scenarios of the classical allosteric models such as Monod-Wyman-Changeux (MWC) (Monod et al. 1965) and Koshland-Nemethy-Filmer (KNF) (Koshland et al. 1966) are successful in describing structured systems, their applicability diminishes

Fig. 3 Glucocorticoid receptor N-terminal domain isoform stability contributes to in vivo activity and refolds to a globular conformation. **a.** Conformational transitions of GR-NTD isoforms (A, C2, and C3) induced by TMAO. Transitions were monitored by tryptophan emission fluorescence intensity at 338 nm with excitation at 295 nm. Emission intensities were normalized to that at 0 M TMAO. Detailed experimental conditions have been previously described (Li et al. 2012). **b.** Relative activities of GR isoforms (Lu and Cidlowski 2005) plotted against the experimentally measured stabilities ($\Delta G_{U \rightarrow F}^0$) from Fig. 3a. The plot indicates that stability correlates to relative activity in a non-linear fashion. **c.** Experimentally measured *m*-values are plotted against the protein length (*N*)—i.e., number of amino acids. Shown are globular proteins (see (Li et al. 2012) for more details) in numbered gray circles: (1) Barstar, (2) RCAM-T1, (3) *P* protein, and (4) Nank 1–7. The GR translational isoforms from Fig. 3a fall on the same line as the globular proteins indicating that they are folding to a conformation that is thermodynamically indistinguishable from a globular fold (Li et al. 2012)

when considering dynamic and disordered systems (Hilser et al. 2012). In particular, neither of the two classical models (KNF or MWC) addresses “how” allosteric sites are coupled to one another (Hilser et al. 2012; Motlagh et al. 2014). To

address this issue and to account for any degree of conformational degeneracy, we can describe the overall energetics in terms of intrinsic energies of conformational change within domains, and coupling energies between them (Hilser and Thompson 2007; Hilser et al. 2012). This can be done by three simple and well-established assumptions: (1) The protein exists in a conformational ensemble in solution, the hallmark of an allosteric protein (Hilser et al. 2012; Koshland et al. 1966; Monod et al. 1965; Motlagh et al. 2014); (2) binding sites for ligands are separated into different domains (Branden and Tooze 1999); and (3) the domains are coupled to one another, i.e., the essence of the allosteric effect (Wyman and Gill 1990). From this simple articulation, it is readily possible to derive a rigorous formalism to quantitatively relate allosteric free energy changes to intrinsic stability changes, interaction energies, and intrinsic binding constants.

Using these assumptions, the simplest allosteric protein can be approximated as two coupled domains: a regulatory effector site (domain I) and an active functional site (domain II) (Fig. 4a). Each of these sites can bind their functional ligand and can exist in their intrinsic conformational ensemble (Fig. 4a, ligands A and B). In the case of an allosteric protein, the two states are high-affinity (H) and low-affinity (L). These states can also be treated as folded (F) and unfolded (U) which is elaborated on in the next section. The important feature is that there is a binding affinity difference between the two conformations. Thus, there are a minimum of four states required to describe such an allosteric protein representing every possible combination of domains I and II being in the H or L states (Fig. 4b, column 1). The most important parameter of this model is the interaction energy between the two domains (Δg_{int} , Fig. 4b, column 3) because it allows the domains to sense one another. This coupling energy is what makes the EAM unique compared to MWC and KNF; it treats the interaction energy as an intrinsic part of the molecule and does not implicitly treat it in macroscopic equilibria. Indeed, as the table shows, the relative energy of each state is comprised of the conformational free energy and the interaction energy:

$$\Delta G_{state} = \Delta G_{conf} + \Delta g_{int} \quad (2)$$

If we choose the fully H (or F) state as our reference state, then there exists an intrinsic energetic cost for changing the conformation of each domain to L (or U) (i.e., what would the cost be in the absence of the other domain) as well as an interaction energy ($\Delta g_{I,II}$) cost associated with making that change in the presence of the other domain (Fig. 4b, columns 2 and 3). In general terms, summing the free energies from conformational transitions and interactions, it is possible to calculate the free energy of each state. From these relative free energies, it is then a straightforward matter to calculate the

statistical weight of each state (S_i) and the partition function (Q), which is the sum of all statistical weights:

$$Q = \sum_{\text{states}, i} S_i = S_{\text{HH}} + S_{\text{LH}} + S_{\text{HL}} + S_{\text{LL}} \quad (3)$$

and the probability ($P_i = S_i/Q$) of any single state or groups of states. For instance, one can determine how the probability of the molecule being active changes upon addition of effector ligand A, by considering the statistical weights of all states that have domain II in the H conformation (i.e., S_{HH} and S_{LH}). We note that in the absence of ligand, the probability is

$$P_{\text{active}}([A] = 0) = \frac{\sum_{\text{active states}, i} S_i}{Q} = \frac{S_{\text{HH}} + S_{\text{LH}}}{Q} \quad (4)$$

Fig. 4 A thermodynamic model of allostery. **a.** Schematic representation of a two-domain allosteric protein consisting of domains I (red) and II (yellow). The two domains are coupled to one another (gray shaded bar) and can each bind their respective ligand (ligands A and B). **b.** Each domain from Fig. 4a can be in a high-affinity (H) or a low-affinity (L) conformation leading to four total states (HH, LH, HL, and LL in column 1) with the low-affinity conformation being represented as a lack of structure. Each state's free energy contribution comes from its conformation (ΔG_{conf} , column 2) and the coupling interaction (Δg_{int} , column 3). Summing these free energies and Boltzmann weighting them (i.e., $S_{\text{state}} = \text{Exp}[-\Delta G_{\text{state}}/(RT)]$ where R is the universal gas constant and T is absolute temperature) leads to the statistical weights (S_i , column 4) of each state. Summing these yields the partition function (Q), which can be used to calculate the probabilities of each state (column 5). **c.** An example of an ensemble that responds allosterically to ligand A. Shown above is the distribution of the ensemble and probability active (i.e., probability that domain II is in its H or active conformation) prior to addition of ligand A, and below is the ensemble after addition of A. There is a ~50 % activation of the ensemble by adding ligand A. The parameters used in this simulation are $\Delta G_I = -2.3$, $\Delta G_{\text{II}} = -0.7$, $\Delta g_{\text{I,II}} = 1.6$, and $\Delta \text{Lig,A} = -3.0$ kcal/mol. Figure adapted with permission (Hilser and Thompson 2007)

Upon addition of effector (ligand A), mass action dictates that ligand A will preferentially bind the H or F state of domain I (Wyman and Gill 1990) which will change the statistical weight of such states by

$$Z_{\text{Lig,A}} = 1 + K_{\text{a,A}}[A] \quad (5)$$

where $Z_{\text{Lig,A}}$ is the contribution to the statistical weight, $K_{\text{a,A}}$ is the association constant for binding ligand A to effector domain I, and $[A]$ is the concentration of ligand A. This will change the partition function to

$$Q([A] > 0) = Z_{\text{Lig,A}}(S_{\text{HH}} + S_{\text{HL}}) + S_{\text{LH}} + S_{\text{LL}} \quad (6)$$

and the probability of being active to

$$P_{\text{active}}([A] > 0) = \frac{\sum_{\text{active states}, i} S_i}{Q([A] > 0)} = \frac{Z_{\text{Lig,A}} * S_{\text{HH}} + S_{\text{LH}}}{Q([A] > 0)} \quad (7)$$

The observed allosteric response is thus due to the difference in probability of these states with and without ligand A:

$$\Delta P_{\text{active}} = P_{\text{active}}([A] > 0) - P_{\text{active}}([A] = 0) \quad (8)$$

An example for a positively coupled system is depicted in Fig. 4c. In the absence of ligand, the ensemble is distributed such that the LL state is the most populated and the molecule is only 28 % active. However, upon addition of ligand A, states HH and HL are stabilized by the ligand binding free energy, and now HH is the dominant state in the ensemble. Ligand A is able to redistribute the ensemble to 78 % active.

The strength of the EAM lies in the fact that it can be divorced from structure. Even though the H conformations are represented as structured relative to L in Fig. 4, the only requirement is that there is a binding affinity difference between H and L. As such, a conformation or ensemble of states that are not well-structured

Fig. 5 The EAM applied to experimental data from GR. **a.** A schematic representation of the A-NTD isoform of GR. The regulatory (R) region is coupled to the functional (F) region as depicted by the gray bar between the domains. R and F are each able to bind their respective ligand (Cofactors A and B, respectively). **b.** Allowing R and F to be folded (F) and unfolded (U) yields four possible states (column 1) similar to Fig. 4b. Each state has a free energy contribution for its conformation (ΔG_{conf} , column 2) and the interaction (Δg_{int} , column 3). Summing these free energies for each state and Boltzmann weighting them yields the statistical weights (S_i , column 4), partition function (Q), and probabilities (column 5). See Fig. 4b and (Hilser and Thompson 2007; Li et al. 2012) for more details. **c.** Plotted is the free energy dependence on TMAO for the F region, R region, interaction energy, and fully unfolded state of GR based on parameters fitted to the EAM. The data were fit to the fluorescence data in Fig. 3a. Assumptions made include that the fully unfolded state UU and the intermediate FU have a fluorescence emission intensity of 0, and that the fully folded state FF and intermediate UF have a fluorescence intensity of 1 (justified by both tryptophans being in the F region). Data were fit using the linear extrapolation method, and fitted parameters are as follows: $\Delta G_{\text{F} \rightarrow \text{U}}^0(\text{F}) = -8.4 \pm 0.5$ kcal/mol, $m(\text{F}) = 7.4 \pm 0.4$ kcal/(mol*M); $\Delta G_{\text{F} \rightarrow \text{U}}^0(\text{R}) = 0$ kcal/mol, $m(\text{R}) = 2.1 \pm 0.4$ kcal/(mol*M); $\Delta g_{\text{int}}^0 = -3.0 \pm 1.0$ kcal/mol, $m_{\text{int}} = 2.0 \pm 0.9$ kcal/(mol*M). See (Li et al. 2012) for more details. **d.** Based on the fitted parameters from Fig. 3a, it is possible to calculate the probability of the F region being folded with a two-state model (shown as data points with C3-NTD in red and A-NTD in yellow). Using the fitted parameters from Fig. 5c, it is possible to predict the probability of the F region being folded (shown as solid lines). The agreement between experiment and fitted parameters lends credence to the EAM

but can bind ligand can be treated just as easily as two ostensibly folded conformations, or even a coupled folding and binding protein such as an IDP (Hilser et al. 2012). The generality of the model is what allows it to describe the full dynamic continuum in Fig. 1 (Motlagh et al. 2014) and makes it applicable to the GR-NTD example from before (Li et al. 2012).

Application of the ensemble allosteric model to glucocorticoid receptor

Because the EAM is articulated in general terms, we can apply it to the specific coupled folding and binding process that the GR-NTD undergoes. We will consider the high-/low-affinity (H/L) equilibria as a simple transition between folded/unfolded states for the A-NTD isoform. Because we know that the A-NTD consists of the R and F region, we can treat it as a two-domain protein that is allosterically coupled (Fig. 5a). Derivation of the partition function is identical to Eqs. 1, 2, 3, 4, 5, 6, 7, and 8 and we already have experimental constraints for some of the variables (Fig. 5b). We know that the stability of the F region is that of C3-NTD which was measured, and we also know that the R region is marginally stable such that both the folded and unfolded exist in solution. Thus, we can treat the stability of the R region as ~ 0 kcal/mol to a first approximation (Li et al. 2012). Sensitivity analyses reveal that the fitted parameters and interpretation do not

change by varying ΔG_{R} significantly (i.e., between -1.5 and 1.5 kcal/mol) (Li et al. 2012).

Using these estimates, we can fit the experimental data to extract the exact value of coupling ($\Delta g_{\text{I,II}}$ which is in this case $\Delta g_{\text{R,F}}$) between the R and F region (Fig. 4b). We find that the fitted value is $\Delta g_{\text{int}} = -3.0$ kcal/mol which indicates that the R and F region are negatively coupled. This comes as no surprise, since addition of the R region to F decreases its stability and in vivo activity (Fig. 3b). This also supports the hypothesis that the coupled folding and binding process is in fact

occurring and we are probing the functional state by increasing the concentration of TMAO.

The strength of a model such as the EAM is that it can be used to make quantitative testable predictions about the ensemble as energetic perturbations are made. For instance, one can ascertain what the TMAO dependence of each state's free energy is by fitting to the fluorescence data (Fig. 5c). From fitting these data, one can then calculate what the probability of the folded state is as a function of TMAO and quantitatively predict the experimental data (Fig. 5d). These osmolyte dependences (Fig. 5c) and predictive capabilities (Fig. 5d) reveal the mechanistic basis of increasing stability upon truncation between the A-NTD and C3-NTD isoform—i.e., the negative coupling between these regions changes the net stability of the whole protein by modulating the intermediates that are populated at elevated concentrations of osmolyte. Taken together, this indicates that ID molecules may tune the stabilities of the folded forms of the molecule by modulating the conformational ensemble.

Conclusion

Allostery is central to the regulation of almost all biological processes. With the realization that proteins do not require stable, unique structures to perform their functions, the need for mechanistic models that can unify allostery in structured, dynamic, and disordered systems has emerged. We have shown previously (Motlagh et al. 2014) that the EAM is not only capable of serving this purpose but can even provide insight to phenomena that are difficult to understand using classic deterministic models of allostery (Koshland et al. 1966; Monod et al. 1965; Hilser et al. 2012). For instance, we have shown that when the entirety of the ensemble is considered, the observed allosteric response is a convolution of competing mechanisms, with some sub-ensembles being responsible for agonist behavior and others being responsible for antagonist behavior with respect to a given allosteric effector. By modulating the relative contributions of the different sub-ensembles (via, for example, the addition of a second allosteric effector), allosteric proteins cannot only change the degree of coupling between two sites, the actual sign of the observed coupling can change, transforming activation (or agonism) into repression (or antagonism), or vice versa. Such experimental observations have been evident in therapeutically relevant cases for decades (Gielen et al. 2005; Hol et al. 1997; Katzenellenbogen et al. 1997), but even a qualitative model for how this occurs has been elusive. The fact that this phenomenon emerges as a straightforward consequence of the heterogeneity in the ensemble suggests that this and other complex phenomena can be understood when the full spectrum of binding and conformational states are considered.

Perhaps the most important aspect of the EAM is that it highlights the probabilistic nature of allostery. Classic views of allostery that focus on an understanding of the “pathway” through the protein often conveys a deterministic or at least homogenous picture of the coupling process, whereby each structural change is coupled in an obligatory fashion to another structural change. Although the mere existence of rheostat-like tunability in allosteric responses (even in structured systems) constitutes *prima facie* evidence that allosteric mechanism cannot be deterministic, the structural-mechanical view of allostery, whereby allostery is reconciled in the context of propagated structural changes between the coupled sites, is still the reigning paradigm. By applying the EAM (or any ensemble model that accounts for the full spectrum of binding and conformational states), it can be seen how the tunability is a direct consequence of changing the stabilities of states, which redistributes the ensemble and modulates the signal in analog fashion. Because nature produces isoforms of proteins (and even RNAs) of different lengths and with different modifications, which will affect the stability and thus redistribute the ensemble, the connection between these modifications, their energetic impact, and the observed allostery is derived directly from the EAM. Of course, it is of little surprise that nature has taken advantage of this ability to tune allosteric mechanism by expressing these isoforms with different *in vivo* activities in different ratios and in different tissues (Lu and Cidlowski 2005).

Acknowledgments We would like to thank Jordan T. White for helpful comments on the manuscript.

Compliance with Ethical Standards

Funding This study was funded by the National Science Foundation grants MCB-0446050 and MCB-1330211 and the National Institutes of Health grants R01-GM063747 and T32-GM008403.

Conflict of interest Hesam N. Motlagh declares that he has no conflict of interest. Jeremy A. Anderson declares that he has no conflict of interest. Jing Li declares that she has no conflict of interest. Vincent J. Hilser declares that he has no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent No individual participants were included in this study.

References

- Adair GS (1926) The oxygen dissociation curve of hemoglobin. *J. Biol. Chem.*
- Auton M, Bolen DW (2005) Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proc Natl Acad Sci U S A* 102: 15065–15068

- Branden C, Tooze J (1999) Introduction to protein structure, 2nd edn. Garland Science, New York
- Changeux JP (2005) Allosteric mechanisms of signal transduction. *Science* 308:1424–1428. doi:[10.1126/science.1108595](https://doi.org/10.1126/science.1108595)
- Cooper A, Dryden DT (1984) Allostery without conformational change. A plausible model. *Eur Biophys J* 11:103–109. **A classic paper that first proposed the concept of dynamic allostery. This has proven to be prescient as many systems have been demonstrated to use dynamics to drive the allosteric response**
- Cui Q, Karplus M (2008) Allostery and cooperativity revisited. *Protein Sci* 17:1295–1307
- Daily MD, Gray JJ (2009) Allosteric communication occurs via networks of tertiary and quaternary motions in proteins. *PLoS Comput Biol* 5, e1000293
- Ferreon ACM, Ferreón JC, Wright PE, Deniz AA (2013) Modulation of allostery by protein intrinsic disorder. *Nature* 498:390–394. **A recent single-molecule study on a disordered viral protein that can switch from positively cooperative to negatively cooperative binding. This is an example of how the ensemble nature of allostery must be considered to reconcile the results**
- Freiburger LA, Baettig OM, Sprules T, Berguis AM, Auclair K, Mittermaier AK (2011) Competing allosteric mechanisms modulate substrate binding in a dimeric enzyme. *Nat Struct Mol Biol* 18:288–294
- Garcia-Pino A et al (2010) Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell* 142:101–111
- Gielen SC, Burger CW, Kuhne LC, Moghaddam P, Blok LJ (2005) Analysis of estrogen agonism and antagonism of tamoxifen, raloxifene, and ICI182780 in endometrial cancer cells: a putative role for the epidermal growth factor receptor ligand amphiregulin. *Soc Gynecol Invest* 12:55–66
- Hilser VJ, Thompson EB (2007) Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc Natl Acad Sci U S A* 104:8311–8315. **This was the first theoretical study to suggest that intrinsic disorder may have roles in allostery. This was also the first articulation of the EAM**
- Hilser VJ, Wrabl JO, Motlagh H (2012) Structural and energetic basis of allostery. *Annu Rev Biophys* 41:585–609
- Hol T, Cox MB, Bryant HU, Draper MW (1997) Selective estrogen receptor modulators and postmenopausal women's health. *J Women's Health* 6:523–531
- Kalodimos CG (2012) Protein function and allostery: a dynamic relationship. *Ann N Y Acad Sci* 1260:81–86
- Katzenellenbogen BS, Montano MM, Ekena K, Herman ME, McNerney ME (1997) Antiestrogens: mechanisms of action and resistance in breast cancer. *Breast Cancer Res Treat* 44:23–38
- Koshland DE Jr, Nemethy G, Filmer D (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* 5:365–385
- Kumar R, Baskakov IV, Srinivasan R, Bolen DW, Lee JC, Thompson EB (1999) Interdomain signaling in a two-domain fragment of the human glucocorticoid receptor. *J Biol Chem* 274:24737–24741
- Kumar R, Lee JC, Bolen DW, Thompson EB (2001) The conformation of the glucocorticoid receptor AF1/tau1 domain induced by osmolyte binds co-regulatory proteins. *J Biol Chem* 276:18146–18152
- Kumar R, Thompson EB (2010) Influence of flanking sequences on signaling between the activation function AF1 and DNA-binding domain of the glucocorticoid receptor. *Arch Biochem Biophys* 496:140–145
- Lavery DN, McEwan IJ (2005) Structure and function of steroid receptor AF1 transactivation domains: induction of active conformations. *Biochem J* 391:449–464. doi:[10.1042/BJ20050872](https://doi.org/10.1042/BJ20050872)
- Li J, Motlagh HN, Chakuroff C, Thompson EB, Hilser VJ (2012) Thermodynamic dissection of the intrinsically disordered N-terminal domain of human glucocorticoid receptor. *J Biol Chem* 287:26777–26787. doi:[10.1074/jbc.M112.355651](https://doi.org/10.1074/jbc.M112.355651)
- Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VU, Dunker AK (2006) Intrinsic disorder in transcription factors. *Biochemistry* 45:6873–6888
- Lu NZ, Cidlowski JA (2005) Translational regulatory mechanisms generate N-terminal glucocorticoid receptor isoforms with unique transcriptional target genes. *Mol Cell* 18:331–342. doi:[10.1016/j.molcel.2005.03.025](https://doi.org/10.1016/j.molcel.2005.03.025)
- Monod J, Wyman J, Changeux JP (1965) On the nature of allosteric transitions: a plausible model. *J Mol Biol* 12:88–118
- Motlagh HN, Wrabl JO, Li J, Hilser VJ (2014) The ensemble nature of allostery. *Nature* 508:331–339
- Muirhead H, Perutz MF (1963) Structure of haemoglobin: a three-dimensional Fourier synthesis of reduced human haemoglobin at 5.5 Å resolution. *Nature* 199:633–638
- Nussinov R, Tsai CJ (2013) Allostery in disease and drug discovery. *Cell* 153:293–305
- Nussinov R, Tsai CJ, Ma B (2013) The underappreciated role of allostery in the cellular network. *Annu Rev Biophys* 42:169–189
- Perutz MF (1970) Stereochemistry of cooperative effects in haemoglobin: haem-haem interaction and the problem of allostery. *Nature* 228:726–739
- Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G, North AC (1960) Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-Å resolution, obtained by x-ray analysis. *Nature* 185:416–422
- Perutz MF, Wilkinson AJ, Paoli M, Dodson GG (1998) The stereochemical mechanism of the cooperative effects in hemoglobin revisited. *Annu Rev Biophys Biomol Struct* 27:1–34
- Popovych N, Sun S, Ebright RH, Kalodimos CG (2006) Dynamically driven protein allostery. *Nat Struct Mol Biol* 13:831–838
- Sevcsik E, Trexler AJ, Dunn JM, Rhoades E (2011) Allostery in a disordered protein: oxidative modifications to α -synuclein act distally to regulate membrane binding. *J Am Chem Soc* 133:7152–7158
- Smock RG, Gierasch LM (2009) Sending signals dynamically. *Science* 324:198–203
- Street TO, Courtemanche N, Barrick D (2008) Protein folding and stability using denaturants. *Methods Cell Biol* 84:295–325
- Tantos A, Han KH, Tompa P (2012) Intrinsic disorder in cell signaling and gene transcription. *Mol Cell Endocrinol* 348:457–465
- Tzeng SR, Kalodimos CG (2009) Dynamic activation of an allosteric regulatory protein. *Nature* 462:368–372. doi:[10.1038/nature08560](https://doi.org/10.1038/nature08560)
- Tzeng SR, Kalodimos CG (2012) Protein activity regulation by conformational entropy. *Nature* 488:236–240
- Wand AJ (2013) The dark energy of proteins comes to light: conformational entropy and its role in protein function revealed by NMR relaxation. *Curr Opin Struct Biol* 23:75–81. doi:[10.1016/j.sbi.2012.11.005](https://doi.org/10.1016/j.sbi.2012.11.005)
- Wyman J, Gill S (1990) Binding and linkage: functional chemistry of biological macromolecules