Published in final edited form as:
Biochem Soc Trans. 2012 October ; 40(5): 975–980. doi:10.1042/BST20120163.

Interplay between allostery and intrinsic disorder in an ensemble

Hesam N. Motlagh[*,1], **Jing Li**[*,†,1], **E. Brad Thompson**[‡,§], and **Vincent J. Hilser**[*,†,2]

[*]T.C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD 21218, U.S.A

[†]Department of Biology, Johns Hopkins University, Baltimore, MD 21218, U.S.A

[‡]Center for Nuclear Receptors and Cell Signaling, Department of Biology and Biochemistry, University of Houston, Houston, TX 77204-5056, U.S.A

[§]Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX 77555-1068, U.S.A

Abstract

Allostery is a biological phenomenon of critical importance in metabolic regulation and cell signalling. The fundamental premise of classical models that describe allostery is that structure mediates ‘action at a distance’. Recently, this paradigm has been challenged by the enrichment of IDPs (intrinsically disordered proteins) or ID (intrinsically disordered) segments in transcription factors and signalling pathways of higher organisms, where an allosteric response from external signals is requisite for regulated function. This observation strongly suggests that IDPs elicit the capacity for finely tunable allosteric regulation. Is there a set of transferable ground rules that reconcile these disparate allosteric phenomena? We focus on findings from the human GR (glucocorticoid receptor) which is a nuclear transcription factor in the SHR (steroid hormone receptor) family. GR contains an intrinsically disordered NTD (N-terminal domain) that is obligatory for transcription activity. Different GR translational isoforms have various lengths of NTD and by studying these isoforms we found that the full-length ID NTD consists of two thermodynamically distinct coupled regions. The data are interpreted in the context of an EAM (ensemble allosteric model) that considers only the intrinsic and measurable energetics of allosteric systems. Expansion of the EAM is able to reconcile the paradox that ligands for SHRs can be agonists and antagonists in a cell-context-dependent manner. These findings suggest a mechanism by which SHRs in particular, and IDPs in general, may have evolved to couple thermodynamically distinct ID segments. The ensemble view of allostery that is illuminated provides organizing principles to unify the description of all allosteric systems and insight into ‘how’ allostery works.

Keywords

allostery; conformational fluctuation; ensemble; glucocorticoid receptor; intrinsic disorder; protein stability

[2]To whom correspondence should be addressed (Hilser@jhu.edu).
[1]These authors contributed equally to this work.

Introduction

Allostery is an essential biological phenomenon playing crucial roles in metabolic regulation and cell signalling, and conferring on cells the ability to differentiate between subtle chemical cues. Arguably the most challenging issue in ascribing a quantitative model for allostery is that it must manifest ‘action at a distance’, i.e. the regulatory site may be several angstroms away from the active site. Is it possible to have a general quantifiable, explanatory and transferable set of rules that describes this phenomenon or are the determinants of allostery unique to each system? Initial proposals to explain allostery in human haemoglobin yielded two influential models: the MWC (Monod–Wyman–Changeux) or ‘concerted’ model [1] and the KNF (Koshland–Nemethy–Filmer) or ‘sequential’ model [2]. Excellent detailed reviews of these historic models are available in [3,4]. During the era when these models were introduced, the first atomic resolution structural models of proteins were being elucidated playing a significant role in the scientific thinking of allostery [5]. What followed was an influential proposal that the cooperative transition in haemoglobin’s quaternary structure was mediated by a mechanical site-to-site perturbation of atoms and side chains [6]. The culmination of these data and success in reproducing some of the phenomenological observations suggested that all allosteric systems could be understood through a structural or single-molecule view of allostery [5].

Previously, emerging theory and accumulating evidence has called into question the structural view of allostery [5,7–14]. Even in human haemoglobin, both the MWC and KNF models fail to reproduce the co-operative binding free energies of oxygen [15,16]. Perhaps the most puzzling reports have been in systems where dynamics and disorder mediate allostery, an observation that is irreconcilable in terms of structure. These well-documented systems include phenomena such as allostery in the absence of a structural pathway [8,11], allostery without conformational change [9], allostery coupled to local unfolding [12,13], allosteric communication from surface mutations that do not affect structure [10,14], allosteric communication in disordered segments [7] and, perhaps most intriguing, the observation that allosteric ligands can act as positive and negative regulators in a cell-context-dependent manner [17–19]. Even more revealing is the realization that IDPs (intrinsically disordered proteins) are enriched in signalling pathways and hyperabundant in transcription factors of higher organisms, suggesting their capacity to finely tune allosteric responses [20]. These observations not only call into question the structural view of allostery, but also undermine the visually appealing notion of the ‘structure–function’ paradigm.

Many of the ID (intrinsically disordered) regions in transcription factors undergo coupled folding upon binding to their co-regulatory partners, suggesting that the functional conformation is a folded state [21,22]. This notion is supported by reports of osmolyte-induced folded states of IDPs showing increased binding affinity to their respective co-regulators [7,23,24].However, whether large ID segments contain thermodynamically and/or functionally distinct regions, the degree of coupling between these ID regions, and how these influence function are poorly characterized and not well understood. In the present paper, we discuss intrinsic disorder and allostery in the human GR (glucocorticoid receptor), a hormone-dependent nuclear transcription factor in the SHR (steroid hormone receptor)

family. Through studying the naturally occurring translational isoforms of GR with various lengths of their ID NTDs (N-terminal domains), we find not only that the full-length GR NTD consists of thermodynamically distinct coupled ID regions, but also that the resultant stabilities of these isoforms correlate with their *in vivo* activity [7]. The data are interpreted in the context of an EAM (ensemble allosteric model) that considers only the experimentally measurable intrinsic energetics of allosteric systems [5,25,26]. An expansion of the EAM is able to reconcile the puzzling observation that certain ligands act allosterically on members in the SHR family as positive and negative regulators in a context-dependent manner [17–19,25]. The culmination of these data suggests a mechanism by which SHRs in particular, and IDPs in general, may have evolved to couple thermodynamically distinct ID segments that are contiguous in sequence. The ensemble view of allostery that is illuminated provides a vehicle to interpret ‘how’ allostery works, possibly in all systems.

Allosteric coupling between ID segments in the NTD of human GR

Because TFs (transcription factors) must respond appropriately in magnitude to external signals, allosteric coupling is critical for proper TF function. The importance of allosteric response in TFs can be appreciated by the wide range of cancers caused by TF dysregulation [27]. Nonetheless, how TFs use structure (or intrinsic disorder) to encode the capacity for tunable allosteric coupling is not well understood.

The SHR family is an ideal target for investigating allostery and its relation to intrinsic disorder because the domain organization is well conserved and most of the members contain a long disordered NTD that is required for proper transcription function and regulation [28–30]. SHRs are hormone-dependent nuclear TFs that play key roles in organ development, metabolite homoeostasis, and stress and inflammatory responses [28]. SHRs typically contain three domains: an ID NTD, a DBD (DNA-binding domain), and an LBD (ligand-binding domain) as depicted in Figure 1(A). The ID NTDs of SHRs are required for transcription activation and regulation through their AF1 (activation function 1) region serving as a hub to recruit co-regulators to form the final transcription complex [F region (functional region) in Figure 1A] [28–30]. Interestingly, the ID NTDs of different SHRs have various lengths and no sequence conservation, yet each one contains an AF1 region in addition to other ID segments. What are the roles of these ID segments outside of AF1 within the NTD of SHRs? In the progesterone receptor and the GR, different translational isoforms vary only in the length of their ID NTD, with each isoform corresponding to a different transcriptional activity [18]. In particular, GR has eight translational isoforms with various activities, different tissue distributions and unique sets of regulated genes [29].Captivatingly, the only difference in the active GR isoforms is the lengths of ID segments with extremely well conserved alternative start sites outside of the AF1 region [7] (Figure 1B). This striking observation strongly suggests that the ID region, which contains multiple translational isoform start sites, serves as a regulatory region for GR function and may consist of thermodynamically coupled regions.

It is well established that IDPs usually undergo coupled folding and binding when they encounter their binding partners, with the folded conformation often serving as the functional state [21,22,31]. It is thus informative to measure the free energy of folding ID

domains because such information reports on the inherent probability of being in the potentially active state. The protective osmolyte TMAO (trimethylamine N-oxide) has been used to induce folding in IDPs to access the free energy difference between the unfolded (inactive) and folded (active) states of the ID NTDs of three human GR translational isoforms A, C2 and C3 [29] (Figure 1B). The experimentally determined stabilities showed a clear relationship with the known *in vivo* transcription activities of their corresponding GR translational isoforms (Figure 1C). This suggests that the stability of the ID NTDs can be used as a mechanism to regulate activity of GR translational isoforms. Interpreting the data within the context of the two-domain EAM reveals that the full-length ID NTD of GR can be divided into two functionally distinct regions: an F region, which contains the co-regulatory binding sites, and an R region (regulatory region), which serves to regulate the stability of the F region and hence GR activity (Figure 1D). The F and R regions are thermodynamically coupled to each other, with unfavourable interaction energy between them, i.e. folding of one domain disfavours folding of the other.

Remarkably, a prediction of the EAM is that two domains are maximally coupled when one or both of the domains are in the unfolded conformation the majority of the time, i.e. ID [25,26]. This suggests that protein chains that have ID regions may actually consist of distinct coupled thermodynamic regions, a notion that is supported by the study presented above [7]. The ideas presented above provide a framework to dissect the thermodynamic basis of allosteric coupling between ID regions in a protein.

The three-domain EAM recapitulates agonism/antagonism switching from allosteric ligands

Since the two-domain EAM described above can be used to explore the role of coupling between two ID segments or one regulatory domain in an allosteric protein, a simple expansion can provide insight into systems with multiple regulatory domains. This is of particular interest to TFs that are essential for proper cell growth, development and differentiation in multicellular organisms [32].Historically, ligands that bind to TFs are classified dichotomously as either positive or negative regulators of the protein they control [32,33]. This notion has been shattered by well-documented observations that ligands for SHRs can elicit both agonistic and antagonistic effects in a cell-context-dependent manner [17–19]. Perhaps most troubling is that this observation is irreconcilable with the classic models of allostery [5].

Expansion of the EAM reconciles this puzzling phenomenon by treating an allosteric protein as three domains that can interact with one another [25] (Figure 2A). By generalizing the folded and unfolded states in the two-domain EAM above to high-affinity relaxed (R) and low-affinity tense (T) states respectively, we can begin to dissect the energetic requisites of agonism/antagonism switching (Figure 2B, first column). The important feature that allows each domain to ‘sense’ the other is that the free energy of each state relative to the native or reference state (i.e. RRR) has a contribution from the transition to the T state (ΔG_i; Figure 2B, second column) in addition to the free energy of breaking the interaction ($\Delta g_{int,i-j}$; Figure 2B, third column) between adjacent R states for the two domains in question:

$$\Delta G_{\text{state}} = \sum_{i,\text{Tense}} \Delta G_i + \sum_{i-j,\text{broken int}} \Delta g_{\text{int},i-j} \quad (1)$$

Boltzmann weighting each state's free energy yields statistical weights (Figure 2B, fourth column), and summing these statistical weights yields the partition function that thermodynamically describes the ensemble:

$$Q = \sum_i e^{-\left(\frac{\Delta G_i}{RT}\right)} = \sum_i S_i \quad (2)$$

where each state *i* summed over in eqn (2) is represented in Figure 2(B). Dividing each state's statistical weight by the partition function yields the probability of being in any conformation (Figure 2B, fifth column):

$$P_{\text{state}j} = \frac{S_j}{Q} = \frac{S_j}{\sum_i S_j} \quad (3)$$

By monitoring changes to the partition function, it becomes readily possible to dissect how the probabilities of states change from energetic perturbations to the ensemble. Because we are concerned with allosteric effects, the energetic perturbation and mechanistic interpretation we consider is ligand binding to domains 1 and 2 as regulatory sites and domain 3 as the functional site, i.e. domain 3 must be in the R conformation for this molecule to be active. Derivation of the equations to describe agonism/antagonism switching is outside the scope of the present mini-review; however, a qualitative grasp can be gleaned from basic linkage principles which dictate that the introduction of ligand will lead to preferential binding to states with the highest affinity [34]. Therefore ligand A or B will bind preferentially to states with domain 1 or 2 in the R conformation, which in turn will increase the bound state's statistical weight by the factor $Z_{\text{Lig,X}}$, i.e.=$(1+K_a[X])$, where K_a is the intrinsic association constant and [X] is the concentration of ligand. Monitoring the probability change of domain 3 in the R conformation (i.e. the active conformation) from ligand A normalized to the free energy of ligand binding yields the CR (coupling response) [5,25,26]:

$$\text{CR}_{3,\text{A}}([\text{B}]=0) = \frac{\Delta P_{3,R}}{\Delta \ln(Z_{lig,\text{A}})} = \frac{P_{3,R}([\text{A}]>0|[\text{B}]=0) - P_{3,R}([\text{A}]=0|[\text{B}]=0)}{\ln(Z_{lig,\text{A}})} \quad (4)$$

$\text{CR}_{3,\text{A}}$ is a measure of the sensitivity of domain 3 to ligand A and the sign of $\text{CR}_{3,\text{A}}$ determines whether ligand A is an agonist (positive) or an antagonist (negative), as this indicates whether the ligand increases or decreases the probability of the molecule being active in response to the energetic perturbation ligand A produces. An analogous equations for $\text{CR}_{3,\text{A}}$ can be derived in the presence of ligand B [25].

We can now pose the question of interest: can ligand A switch from being an agonist to an antagonist (or vice versa) if a second ligand B is present at a physically distinct site? Remarkably, the answer is yes, and the set of energetic parameters that describe the

ensemble are robustly encoded and highly degenerate [25]. It should be noted that the interactions between the protein and ligand do not differ in these ensembles that are 'switching-competent' and that the effect is encoded in the thermodynamic architecture. Shown in Figure 2(C) is an example of one such ensemble. The *z*-axis represents the activity change from introduction of ligand A only. Of note is that for a defined stability of domain 1, there exists both a positive and negative response depending on the stability of domain 2 (Figure 2C, yellow circle 1 and 2 respectively). Indeed, all that is required to change the agonistic (positive) response to an antagonistic (negative) response is an energetic perturbation on domain 2. Remarkably, this energetic perturbation can be facilitated easily by physiologically relevant ligand concentrations and free energies of binding [25]. How general is this result?

An exhaustive search of parameter space reveals that the ability to be agonist/antagonist switch competent does not rely on the individual stabilities of domains. Instead, the effect is due to the sign of the interaction energies that poise the ensemble in an energetic 'tug of war'. Since each of the three interaction parameters can be positive or negative, there are eight possible combinations of positive and negative couplings. Interestingly, half are switch-competent (Figure 2D), whereas the other half are committed agonists or antagonists (Figure 2E). The basis of the energetic competition can be revealed if one considers that stabilization of domain 1 by ligand A will have the net effect of agonistically affecting adjacent domains to which it is positively coupled (Figures 2D and 2E, blue arrows) or antagonistically affecting if it is negatively coupled (Figures 2D and 2E, red arrows). This can be appreciated by inspection of Figure 2(D) node 1 denoted + + − (i.e. $\Delta g_{\text{int},1-2}>0$, $\Delta g_{\text{int},2-3}>0$ and $\Delta g_{\text{int},1-3}<0$; similar rules apply to all other architectures). Binding of ligand A will destabilize domain 3 through the negative interaction parameter (Figure 2D, direct effect). However, since the two remaining interactions are positive, domain 3 also senses a stabilizing effect through domain 3 (Figure 2D, indirect effect). Therefore the binding of ligand A has competing effects propagating to domain 3, and the net response that prevails can be modulated by attenuating the stability of domain 2 through ligand binding as in Figure 2C. In contrast, the switch-incompetent architectures do not result in an energetic competition; the net effect is either committed agonism (Figure 2E, left-hand side) or committed antagonism (Figure 2E, right-hand side).

A striking corollary of the EAM can is revealed if the parameter combinations are recast as the probability of the regulatory domains (1 and 2) being in the T conformation in the absence of ligand. The ability to maximally couple two domains, or to be agonist/antagonist switch-competent, is maximized when one or both of the regulatory domains are in the T state the majority of the time [5,25,26]. Recalling that in the limit of the protein folding reaction, the R and T states represent the folded and unfolded states; this implies that IDPs confer the ability to maximally couple domains and also be agonist/antagonist switch-competent. Put another way, this result suggests that one stretch of continuous ID polypeptide chain may be actually two thermodynamically distinct domains coupled to one another. This is in striking agreement with the differential activity of GR NTD translational isoforms [7,18] and supports the notion that the enrichment of ID segments or proteins in cell signalling and TFs confers the ability to finely tune allosteric responses [5,25,26].

Conclusions

Allostery has classically been rationalized in terms of structural changes that can be ascertained from inspection and comparison of the high-resolution structures of the pre- and post-activated complexes. The recognition that many allosteric proteins contain significant sequence segments that are ID, suggests that purely structural interpretations are not likely to illuminate allosteric mechanism in all systems. We propose that an ensemble-based thermodynamic description of allostery (i.e. the EAM) can add clarity to these cases. Indeed, in the case of the ID NTD of GR, the thermodynamic coupling between different ID segments is a critical prediction of the EAM, demonstrating that allostery can be mechanistically understood in the absence of a unique structural context [5,25,26]. Because the EAM is cast in terms of the intrinsic energetics of the co-operative elements within the protein, the binding affinities of the different conformational states of each element, and their interaction energies (all of which are measurable), the EAM provides a vehicle for interpreting allosteric phenomena such as allostery without structural change [10,14], allostery in the absence of structural pathways [8,11] and allosteric communication between ID segments [7]. Furthermore, because allostery is an effect that is based on the distribution of states in the ensemble, and the structural and functional properties of those states, the EAM provides a framework to quantitatively unify the emerging phenomenon of ‘dynamic allostery’ with classic structural models [9,35].

Acknowledgments

Funding

This work was supported by the National Science Foundation [grant number MCB-0446050] and the National Institutes of Health [grant number T32-GM008403].

Abbreviations used

AF1	activation function 1
CR	coupling response
DBD	DNA-binding domain
EAM	ensemble allosteric model
F region	functional region
GR	glucocorticoid receptor
ID	intrinsically disordered
IDP	intrinsically disordered protein
KNF	Koshland–Nemethy–Filmer
LBD	ligand-binding domain
MWC	Monod–Wyman–Changeux
NTD	N-terminal domain

R	relaxed
R region	regulatory region
SHR	steroid hormone receptor
T	tense
TF	transcription factor
TMAO	trimethylamine N-oxide.

References

1. Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: a plausible model. J. Mol. Biol. 1965; 12:88–118. [PubMed: 14343300]
2. Koshland DE, Nemethy G, Filmer D. Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry. 1966; 5:365–385. [PubMed: 5938952]
3. Changeux JP. Allostery and the Monod–Wyman–Changeux model after 50 years. Annu. Rev. Biophys. 2012; 41:103–133. [PubMed: 22224598]
4. Cui Q, Karplus M. Allostery and cooperativity revisited. Protein Sci. 2008; 17:1295–1307. [PubMed: 18560010]
5. Hilser VJ, Wrabl JO, Motlagh HN. Structural and energetic basis of allostery. Annu. Rev. Biophys. 2012; 41:585–609. [PubMed: 22577828]
6. Perutz MF. Stereochemistry of cooperative effects in haemoglobin. Nature. 1970; 228:726–734. [PubMed: 5528785]
7. Li J, Motlagh HN, Chakuroff C, Thompson EB, Hilser VJ. Thermodynamic dissection of the intrinsically disordered N-terminal domain of human glucocorticoid receptor. J. Biol. Chem. 2012; 287:26777–26787. [PubMed: 22669939]
8. Daily MD, Upadhyaha TJ, Gray JJ. Contact rearrangements form coupled networks from local motions in allosteric proteins. Proteins: Struct., Funct., Bioinf. 2008; 71:455–466.
9. Popovych N, Sun S, Ebright RH, Kalodimos CG. Dynamically driven protein allostery. Nat. Struct. Mol. Biol. 2006; 13:831–838. [PubMed: 16906160]
10. Schrank TP, Bolen DW, Hilser VJ. Rational modulation of conformational fluctuations in adenylate kinase reveals a local unfolding mechanism for allostery and functional adaptation in proteins. Proc. Natl. Acad. Sci. U.S.A. 2009; 106:16984–16989. [PubMed: 19805185]
11. Petit CM, Zhang J, Sapienza PJ, Fuentes EJ, Lee AL. Hidden dynamic allostery in a PDZ domain. Proc. Natl. Acad. Sci. U.S.A. 2009; 106:18249–18254. [PubMed: 19828436]
12. Reichheld SE, Yu Z, Davidson AR. The induction of folding cooperativity by ligand binding drives the allosteric response of tetracycline repressor. Proc. Natl. Acad. Sci. U.S.A. 2009; 106:22263–22268. [PubMed: 20080791]
13. Freiburger LA, Baettig OM, Sprules T, Berguis AM, Auclair K, Mittermaier AK. Competing allosteric mechanisms modulate substrate binding in a dimeric enzyme. Nat. Struct. Mol. Biol. 2011; 18:288–294. [PubMed: 21278754]
14. Schrank TP, Elam WA, Li J, Hilser VJ. Strategies for the thermodynamic characterization of linked binding/local folding reactions within the native state: application to the LID domain of adenylate kinase from *Escherichia coli*. Methods Enzymol. 2011; 492:253–282. [PubMed: 21333795]
15. Smith FR, Ackers GK. Experimental resolution of cooperative free energies for the ten ligation states of human hemoglobin. Proc. Natl. Acad. Sci. U.S.A. 1985; 82:5347–5351. [PubMed: 3860865]
16. Gill SJ, Robert CH, Coletta M, Di Cera E, Brunori M. Cooperative free energies for nested allosteric models as applied to human hemoglobin. Biophys. J. 1986; 50:747–752. [PubMed: 3779009]

17. Katzenellenbogen BS, Montano MM, Ekena K, Herman ME, McInerney ME. Antiestrogens: mechanisms of action and resistance in breast cancer. Breast Cancer Res. Treat. 1997; 44:23–38. [PubMed: 9164675]
18. Lu NZ, Cidlowski JA. Translational regulatory mechanisms generate N-Terminal glucocorticoid receptor isoforms with unique transcriptional target genes. Mol. Cell. 2005; 18:331–342. [PubMed: 15866175]
19. Hol T, Cox MB, Bryant HU, Draper MW. Selective estrogen receptor modulators and postmenopausal women's health. J. Women's Health. 1997; 6:523–531. [PubMed: 9356975]
20. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VU, Dunker AK. Intrinsic disorder in transcription factors. Biochemistry. 2006; 45:6873–6888. [PubMed: 16734424]
21. Tantos A, Han KH, Tompa P. Intrinsic disorder in cell signaling and gene transcription. Mol. Cell. Endocrinol. 2012; 348:457–465. [PubMed: 21782886]
22. Uversky VN. Intrinsically disordered proteins from A to Z. Int. J. Biochem. Cell Biol. 2011; 43:1090–1103. [PubMed: 21501695]
23. Baskakov I, Bolen DW. Forcing thermodynamically unfolded proteins to fold. J. Biol. Chem. 1998; 273:4831–4834. [PubMed: 9478922]
24. Kumar R, Lee JC, Bolen DW, Thompson EB. The conformation of the glucocorticoid receptor AF1/tau1 domain induced by osmolyte binds co-regulatory proteins. J. Biol. Chem. 2001; 276:18146–18152. [PubMed: 11279138]
25. Motlagh HN, Hilser VJ. Agonism/antagonism switching in allosteric ensembles. Proc. Natl. Acad. Sci. U.S.A. 2012; 109:4134–4139. [PubMed: 22388747]
26. Hilser VJ, Thompson EB. Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. Proc. Natl. Acad. Sci. U.S.A. 2007; 104:8311–8315. [PubMed: 17494761]
27. Ahmad A, Kumar R. Steroid hormone receptors in cancer development: a target for cancer therapeutics. Cancer Lett. 2011; 300:1–9. [PubMed: 20926181]
28. Griekspoor A, Zwart W, Neefjes J, Michalides R. Visualizing the action of steroid hormone receptors in living cells. Nucl. Recept. Signaling. 2007; 5:e003.
29. Lavery DN, McEwan IJ. Structure and function of steroid receptor AF1 transactivation domains: induction of active conformations. Biochem. J. 2005; 391:449–464. [PubMed: 16238547]
30. McEwan IJ, Lavery D, Fischer K, Watt K. Natural disordered sequences in the amino terminal domain of nuclear receptors: lessons from the androgen and glucocorticoid receptors. Nucl. Recept. Signaling. 2007; 5:e001.
31. Smock RG, Gierasch LM. Sending signals dynamically. Science. 2009; 324:198–203. [PubMed: 19359576]
32. Latchman DS. Transcription factors: an overview. Int. J. Biochem. Cell Biol. 1997; 29:1305–1312. [PubMed: 9570129]
33. Alberts, B.; Johnson, AD.; Lewis, J.; Raff, M.; Roberts, K.; Walter, W., editors. Garland Science. New York: 2007. Molecular Biology of the Cell.
34. Wyman, J.; Gill, SJ. Binding and Linkage: Functional Chemistry of Biological Macromolecules. Mill Valley: University Science Books; 1990.
35. Tzeng SR, Kalodimos CG. Dynamic activation of an allosteric regulatory protein. Nature. 2009; 462:368–372. [PubMed: 19924217]
36. Pei J, Grishin NV. PROMALS: towards accurate multiple sequence alignments of distantly related proteins. Bioinformatics. 2007; 23:802–808. [PubMed: 17267437]

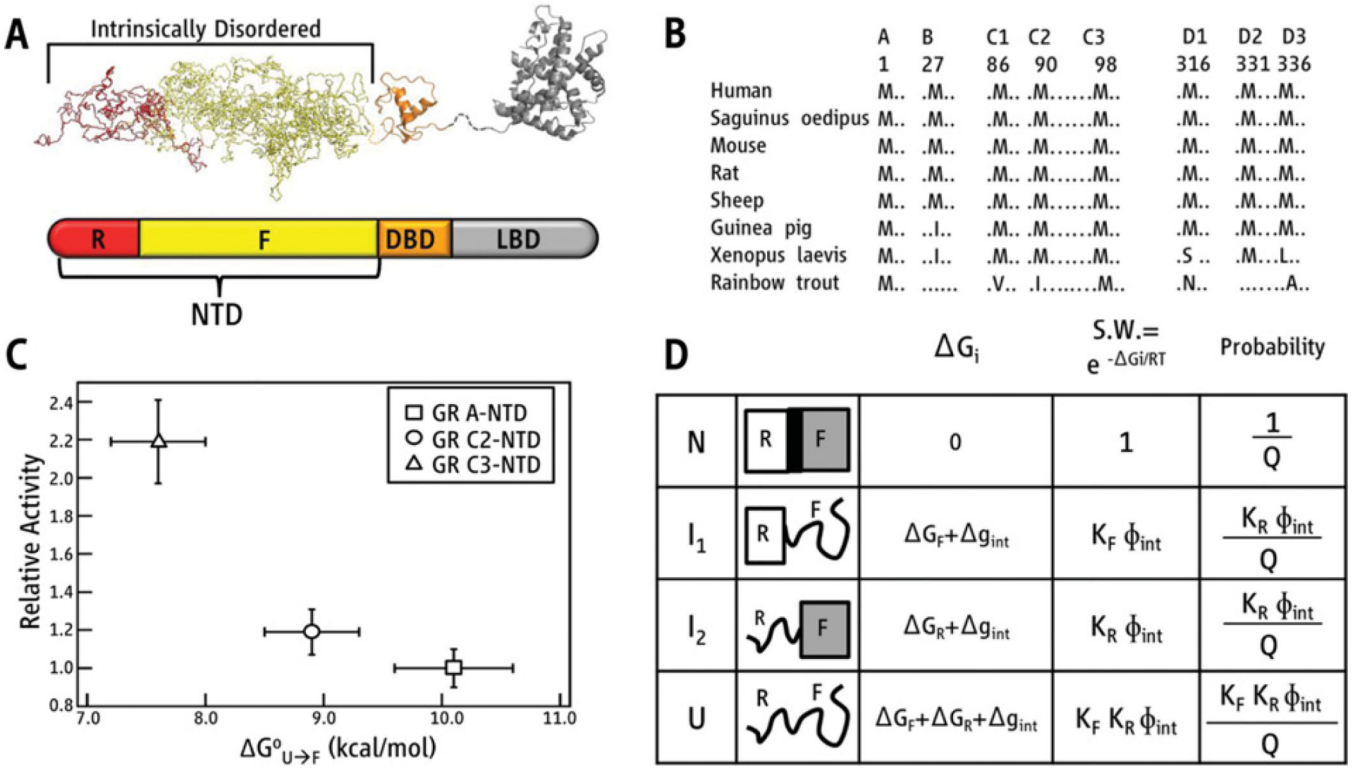
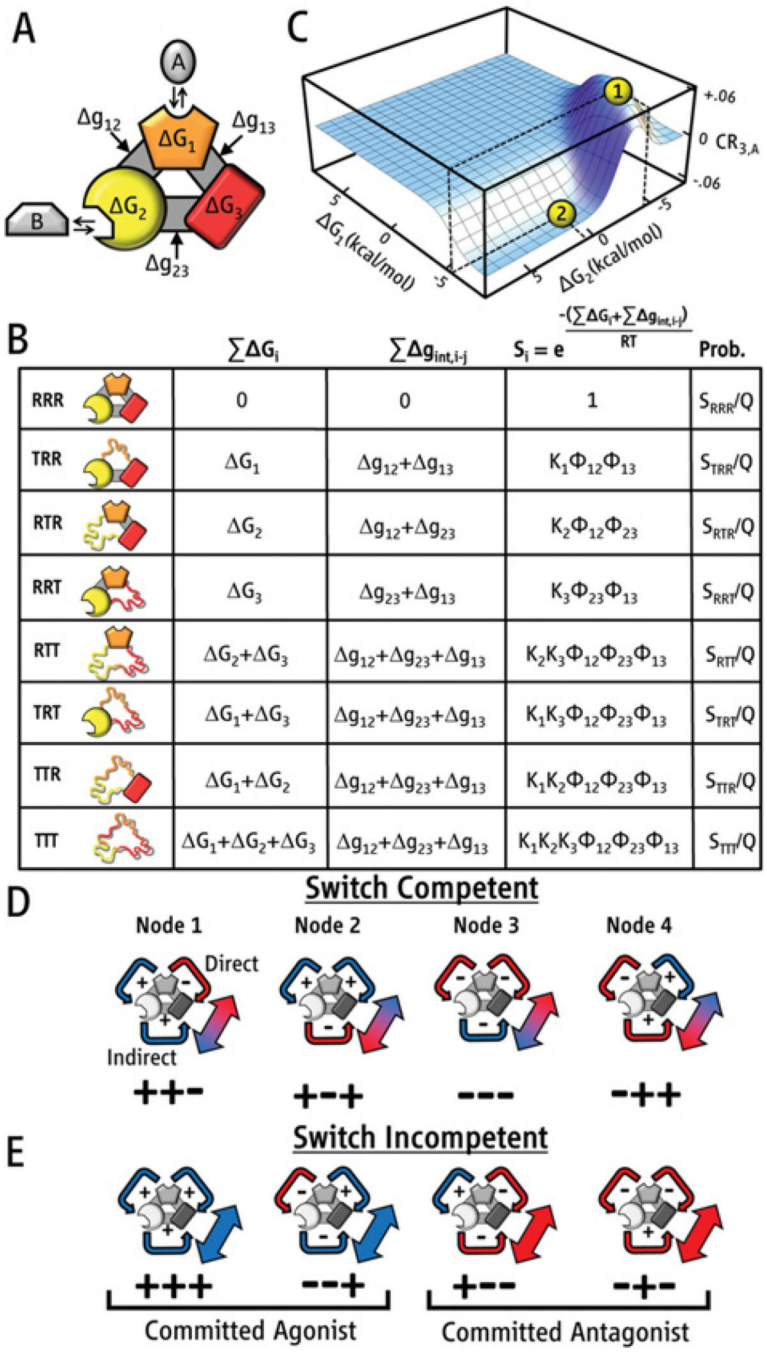


Figure 1. Conserved alternative translational start sites modulate stability and correlate with *in vivo* activity

(**A**) Typical domain architecture of an SHR in general, and GR in particular. The domains are, from N-terminus to C-terminus, denoted NTD, DBD and LBD. The ID NTD is subdivided into the R region (amino acids 1–97 in **B**) and F region (amino acids 98–420 in **B**). (**B**) Conservation of alternative translational start site methionine residues (M) in GR NTD from different species. Alignment was performed using PROMALS3D [36], and amino acids are numbered on top with respect to human GR. The translational isoforms are named A, B, C1, C2, C3, D1, D2 and D3 according to [18]. (**C**) Measured stabilities ($\Delta G^0_{U\rightarrow F}$) of GR translational isoforms ID NTD from TMAO refolding experiments [7] against measured relative activities reported in [18]. Although the relationship is non-linear, the plot indicates that isoforms with more stable NTDs (i.e. further left on the plot) have higher *in vivo* activities. (**D**) Schematic representation of an ensemble of states for a two-domain allosteric protein. The model is presented in the context of the full-length GR NTD containing the R region and F region that are negatively coupled to each other ($\Delta g_{int} < 0$). Since each domain (R or F region) can be folded or unfolded, there are four possible states (i.e. N, I1, I2 and U) with the free energies relative to the native (N) state having contributions from unfolding (ΔG_i) and breaking the interaction (Δg_{int}) [26]. S.W. is the statistical weight of each state. Q is the partition function, which is the sum of the statistical weights of all of the states in the ensemble. K_F and K_R are equilibrium constants between the folded and unfolded state for the F and R region respectively, i.e. $K_F = \exp(-\Delta G_F/RT)$ and $K_R = \exp(-\Delta G_R/RT)$. Φ_{int} is the statistical weight of the interaction between the R and F region, $\Phi_{int} = \exp(-\Delta g_{int}/RT)$.



	$\sum\Delta G_i$	$\sum\Delta g_{int,i-j}$	$S_i = e^{\frac{-(\sum\Delta G_i+\sum\Delta g_{int,i-j})}{RT}}$	Prob.
RRR	0	0	1	S_{RRR}/Q
TRR	ΔG_1	$\Delta g_{12}+\Delta g_{13}$	$K_1\Phi_{12}\Phi_{13}$	S_{TRR}/Q
RTR	ΔG_2	$\Delta g_{12}+\Delta g_{23}$	$K_2\Phi_{12}\Phi_{23}$	S_{RTR}/Q
RRT	ΔG_3	$\Delta g_{23}+\Delta g_{13}$	$K_3\Phi_{23}\Phi_{13}$	S_{RRT}/Q
RTT	$\Delta G_2+\Delta G_3$	$\Delta g_{12}+\Delta g_{23}+\Delta g_{13}$	$K_2K_3\Phi_{12}\Phi_{23}\Phi_{13}$	S_{RTT}/Q
TRT	$\Delta G_1+\Delta G_3$	$\Delta g_{12}+\Delta g_{23}+\Delta g_{13}$	$K_1K_3\Phi_{12}\Phi_{23}\Phi_{13}$	S_{TRT}/Q
TTR	$\Delta G_1+\Delta G_2$	$\Delta g_{12}+\Delta g_{23}+\Delta g_{13}$	$K_1K_2\Phi_{12}\Phi_{23}\Phi_{13}$	S_{TTR}/Q
TTT	$\Delta G_1+\Delta G_2+\Delta G_3$	$\Delta g_{12}+\Delta g_{23}+\Delta g_{13}$	$K_1K_2K_3\Phi_{12}\Phi_{23}\Phi_{13}$	S_{TTT}/Q

Figure 2. Three-domain EAM can recapitulate agonism/antagonism switching

(**A**) Schematic representation of a three-domain allosteric protein composed of domains 1, 2 and 3 with intrinsic stabilities ΔG_1, ΔG_2 and ΔG_3 respectively. Domains 1 and 2 are the regulatory sites with their allosteric effectors represented as ligand A and B respectively. Interactions between the three domains are shown as the grey bars connecting each domain and denoted by Δg_{ij} where *ij* represents the two domains interacting. (**B**) States in the ensemble are represented in the far left-hand column along with their free energy contributions (see the text), Boltzmann statistical weights [note: $K_i = \exp(-\Delta G_i/RT)$ and Φ_{ij}

=exp($-\Delta g_{\mathrm{int},ij}/RT$)] and probabilities of being populated. Each state is denoted by a three-letter code representing whether domains 1, 2 and 3 are in the R or T conformation respectively. (**C**) Example of an agonist/antagonist switching thermodynamic architecture: $\Delta G_1 = -6.8$, $\Delta G_{2,\mathrm{B}=0}(1) = -4.4$, $\Delta G_{2,\mathrm{B}>0}(2) = 0.6$, $\Delta G_3 = -2.7$, $\Delta g_{12} = 6.8$, $\Delta g_{23} = 4.8$, $\Delta g_{13} = -1.9$ and $\Delta g_{\mathrm{Lig,A}} = -5.0$ kcal/mol (1 kcal=4.184 kJ). The CR of domain 3 from ligand A ($\mathrm{CR}_{3,\mathrm{A}}$) is dependent on the stability of domain 1 and 2. Point 1 represents the agonistic response and point 2 represents the antagonistic response at the local maximum and minimum respectively. (**D**) Representation of the interdomain couplings for each of the switch-competent nodes numbered 1–4 represented by three + or − (see the text). Blue or red arrows designate an agonistic or antagonistic net effect on domain 3 respectively. (**E**) Representation of the interdomain couplings for each of the switch-incompetent nodes. The colouring scheme of the arrows is identical with those in (**D**) and shows that certain architectures are committed agonists (left: + + +, − − +) or committed antagonists (right: + − −, − + −), since the net redistribution effect is additive instead of opposing as in (**D**). Adapted from Motlagh, H.N. and Hilser, V.J. (2012) Agonism/antagonism switching in allosteric ensembles. Proc. Natl. Acad. Sci. U.S.A. **109**, 4134–4139 with permission.