AIM for Allostery: Using the Ising Model to Understand Information Processing and Transmission in Allosteric Biomolecular Systems

Michael V. LeVine[1] and **Harel Weinstein**[1,2,*]

Michael V. LeVine: mil2037@med.cornell.edu

[1]Department of Physiology and Biophysics, Weill Cornell Medical College, Cornell University, New York, NY 10065, USA

[2]HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute of Computational Biomedicine, Weill Cornell Medical College, Cornell University, New York, NY 10065, USA

Abstract

In performing their biological functions, molecular machines must process and transmit information with high fidelity. Information transmission requires dynamic coupling between the conformations of discrete structural components within the protein positioned far from one another on the molecular scale. This type of biomolecular "action at a distance" is termed *allostery*. Although allostery is ubiquitous in biological regulation and signal transduction, its treatment in theoretical models has mostly eschewed quantitative descriptions involving the system's underlying structural components and their interactions. Here, we show how Ising models can be used to formulate an approach to allostery in a structural context of interactions between the constitutive components by building simple allosteric constructs we termed *Allosteric Ising Models* (AIMs). We introduce the use of AIMs in analytical and numerical calculations that relate thermodynamic descriptions of allostery to the structural context, and then show that many fundamental properties of allostery, such as the multiplicative property of parallel allosteric channels, are revealed from the analysis of such models. The power of exploring mechanistic structural models of allosteric function in more complex systems by using AIMs is demonstrated by building a model of allosteric signaling for an experimentally well-characterized asymmetric homodimer of the dopamine D2 receptor.

Keywords

allostery; biophysics; Ising model; statistical mechanics; signal transduction; information theory; G protein coupled receptors (GPCRs); dopamine D2 receptor; functional selectivity

*Author to whom correspondence should be addressed; haw2002@med.cornell.edu.

1. Introduction

Complex molecular assemblies and networks of interacting biomolecules mediate many cellular processes, such as cell growth, metabolism, and signaling. Molecular components of such assemblies and networks have been visualized and structurally elucidated at atomic-level resolution with experimental techniques including x-ray crystallography [1], nuclear magnetic resonance (NMR) [2], and cryo-electron microscopy (cryo-EM) [3]. The combination of structure elucidation with the application of biophysical methods reporting on the dynamic properties of the molecules (e.g., single molecule Forster resonance energy transfer (smFRET) [4], electron paramagnetic resonance (EPR) [5], Molecular Dynamics (MD) simulations [6], and elastic network models [7]) has produced detailed information regarding functional mechanisms. The application of these powerful methods of molecular biophysics has illuminated, especially in proteins, the large ensemble of conformations involved in the functional mechanisms of biomolecules, and hence the importance of conformational entropy. This conformational entropy is much higher than expected from crystal structures alone, and the relatively discrete structural elements comprising these systems (*i.e.*, loops, α-helices, β-strands, and a large number of tertiary structures in proteins) often exhibit coupled conformational dynamics. These coupled dynamics are especially crucial in receptor proteins, which are used to process and transmit information in their signaling function. For example, transmembrane receptor proteins, such as the G protein coupled receptors (GPCRs), bind extracellular ligands that trigger receptor "activation", which is reflected by a change in conformation on the intracellular side of the protein where the transduction of the signal into the cell is accomplished [8]. This type of "action-at-a-distance" in the modulation of a specific function is referred to as allostery [9]. While allostery has been documented in many systems and has been suggested to be present in nearly all proteins [10], it is still unclear how most allosteric mechanisms work at the molecular level. A strong theoretical basis for allostery is needed, however, because such mechanisms are ubiquitous and essential for the transduction of signals and transmitting information both within proteins and throughout cellular systems. In addition, while there has been some success in engineering allosteric proteins from pre-existing components and scaffolds, a lack of detailed understanding has placed *de novo* design out of reach [11].

Considerations of theoretical models of allostery have generally followed a thermodynamic approach [9,12,13]. When biochemical measurements of the functional output of proteins can be made, the allosteric efficacy [14], which has also been called the allosteric coupling constant [15], can be used as a good measure of a ligand's allosteric influence on the protein's functional state. For the case of receptors, this downstream signal transduction can be measured experimentally. Assuming that the receptor has two states, *on* and *off*, an allosteric efficacy, α, can be defined as:

$$\alpha = \frac{K_{bound}}{K_{unbound}} \quad (1)$$

where K_{bound} and $K_{unbound}$ are the equilibrium constants for the activation reactions of the receptor when bound or unbound to the allosteric ligand, respectively. An equilibrium constant can be defined in terms of concentrations or rate constants:

$$K = \frac{[R_{on}]}{[R_{off}]} = \frac{k_{on}}{k_{off}} \quad (2)$$

where $[R_{on}]$ and $[R_{off}]$ are the steady state concentrations of the receptor in the *on* and *off* state, respectively, and k_{on} and k_{off} are the corresponding rate constants for the transition to the *on* and *off* states (see Figure 1). The concentrations of the two receptor populations can be inferred from biochemical measurements of function, and the allosteric efficacy of the ligand of interest can be calculated from (1) and (2). When $\alpha > 1$, the *on* state of the receptor is preferred in the presence of ligand and the ligand is considered an agonist (activator of function), and when $\alpha < 1$, the *off* state of the receptor is preferred in the presence of ligand and the ligand is considered an inverse agonist (inhibitor of function). When α is 1, the ligand has no effect on the functional state of the receptor and the ligand is considered a neutral antagonist (inhibitor of activation by another ligand). This type of allostery, in which the equilibrium constant is modified by the ligand, is often described as "K-type", as opposed to those that change enzyme catalysis in terms of k_{cat} or V_{max}, which are described as "V-type" [15].

It is possible to conceptualize the allosteric efficacy of a ligand as a steady state signal-to-noise ratio, where the signal for the presence of a ligand in the binding site is encoded in the receptor on/off equilibrium constant that is sensed by the intracellular proteins that detect the signal by interacting with the receptor population. In the absence of ligand the equilibrium constant is non-zero (*i.e.*, the probability of the receptor being active is non-zero), creating noise.

To obtain a formal definition of the allosteric efficacy in this context, it is possible to write the signal-to-noise ratio, SNR, as:

$$SNR = \frac{P_{signal}}{P_{noise}} \quad (3)$$

where P_x is the power of x defined as:

$$P_x = \frac{1}{T} \int_0^T x(t)^2 dt \quad (4)$$

Therefore, the power of the equilibrium constant signal can be written as:

$$P_K = \frac{1}{T} \int_0^T K(t)^2 dt \quad (5)$$

and because at steady state the equilibrium constant is invariant with time by definition:

$$P_K = K^2 \quad (6)$$

then:

$$\alpha = \sqrt{\frac{P_{K_{bound}}}{P_{K_{unbound}}}} \quad (7)$$

Accordingly, the allosteric efficacy of an agonist is a measure of the signal-to-noise ratio of signaling through the receptor with that agonist. If the ligand is an inverse agonist, the pertinent measure is the equilibrium constant for the inactivation reaction, so that the signal-to-noise ratio is simply α^{-1}. When both the signal and noise are Gaussian, the Shannon-Hartley theorem [16,17] relates the signal-to-noise ratio to the information theoretical channel capacity C (which is the upper limit on the information rate or mutual information), by:

$$C = B \log(1 + SNR) \quad (8)$$

where B is the bandwidth of the channel. While Equation (8) is not directly applicable to the allosteric efficacy, as the signal and noise are not Gaussian, the treatment of allostery as an information transmission process has had much success recently [18–20], and we will confirm a strong relationship between the mutual information and allosteric efficacy later in the manuscript.

An energy-based expression of the allosteric efficacy can be written as the difference in free energies, G, of the four respective states:

$$-RT\log(\alpha) = (G_{on,bound} - G_{on,unbound}) + (G_{off,unbound} - G_{off,bound}) \quad (9)$$

where R is the gas constant and T is the temperature. This model can be extended to systems with multiple ligand binding sites and/or allosterically regulated sites (for a detailed review, see [13]), but it clearly provides only a phenomenological explanation of allostery. According to this description, often considered "the thermodynamic" perspective, allostery occurs because of the differences in free energy of the respective states. However, this conclusion appears to be a definition, *i.e.*, that allostery is the phenomenon in which the stability of the *on* state relative to the *off* state is greater when the ligand is bound, and lesser when the ligand is unbound. From a "structural" perspective, one needs to consider the differences in free energy as emerging from some feature of the underlying network of interacting structural components, and it is this feature that makes the system allosteric.

To understand allostery at a level that explains the structural context for how allosteric biomolecular systems work requires a quantitative theoretical description that bridges the features of the structural components and their interactions, to the thermodynamic allosteric parameters. We address this problem in the next section.

2. Results and Discussion

2.1. The Thermodynamic Allosteric Efficacy as a Function of Local Interactions

We approach the problem of "how allostery works" by studying the statistical mechanics of interacting structural components. These structural components may be any subset of a biomolecular system that can be treated as a unit when described at some level of coarse-graining (*i.e.*, a helix, a β strand, a helical bundle, a binding site, *etc*). The approach we will pursue is conceptually similar to the ensemble allosteric model (EAM) [12], but with the goal of introducing a structural context that can be analyzed analytically. Defining an n-component system X where for a single configuration each component can be in one of an arbitrary number of discrete states, we write the potential energy function of a given configuration of X, U(X), as:

$$U(X) = \sum_{i=1}^{n} U^{conf}(X_i) + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{U^{int}(X_i, X_j)}{2} \quad (10)$$

The first term in (10) represents the conformational energy of each state of each component independent of other components, and the second term represents the pairwise interaction energy between components; all interaction terms when $i = j$ are 0. We can write the probability of any conformation of the system according to the Boltzman distribution as:

$$p(X) = \frac{e^{-\beta U(X)}}{Z} \quad (11)$$

β is $1/k_B T$, where k_B is the Boltzmann constant and T is the temperature in Kelvin. The numerator is known as the Boltzmann factor, and Z is the partition function, which sums over the Boltzmann factors of all states and normalizes the probability:

$$Z = \sum e^{-\beta U(X)} \quad (12)$$

We can then define the specific case of ligand binding to a two-state receptor. This system can be defined as a two-component system in which each component is two-state: one component representing the receptor, R, with states *on* and *off*, and the second component representing the ligand, L, with states *bound* and *unbound*. It should be noted that for the ligand, the conformational energy term represents the component of the binding energy that is independent of the state of the receptor. Using the explicit definition of the concentration:

$$[X] = \frac{N_x}{V} \quad (13)$$

where N_x is the number of molecules of X and V is the volume, we can rewrite (2) with the explicit definition of protein concentration:

$$K=\frac{\frac{Nf_{on}}{V}}{\frac{Nf_{off}}{V}}=\frac{f_{on}}{f_{off}} \quad (14)$$

where N is the total number of receptors and f_{on} and f_{off} are the fraction of receptors in the *on* and *off* states, respectively. Given that the system is ergodic, the frequency of a given state at steady state will converge to the ensemble probabilities. Rewriting (1) by substituting thermodynamic equilibrium constants with ratios of probabilities, we can define the allosteric efficacy as:

$$\alpha \frac{p\left(L=unbound, R=on\right)}{p\left(L=unbound, R=off\right)}=\frac{p\left(L=bound, R=on\right)}{p\left(L=bound, R=off\right)} \quad (15)$$

Using (10) and (11), we can write (15) as:

$$\alpha \frac{e^{-\beta\left[U^{conf}(L=unbound)+U^{conf}(R=on)+U^{int}(L=unbound,R=on)\right]}}{e^{-\beta\left[U^{conf}(L=unbound)+U^{conf}(R=off)+U^{int}(L=unbound,R=off)\right]}}=\frac{e^{-\beta\left[U^{conf}(L=bound)+U^{conf}(R=on)+U^{int}(L=bound,R=on)\right]}}{e^{-\beta\left[U^{conf}(L=bound)+U^{conf}(R=off)+U^{int}(L=bound,R=off)\right]}} \quad (16)$$

Equation (16) reduces to:

$$\alpha=e^{-\beta\left[\left(U^{int}(L=bound,R=on)-U^{int}(L=bound,R=off)\right)+\left(U^{int}(L=unbound,R=off)-U^{int}(L=unbound,R=on)\right)\right]} \quad (17)$$

We then find the analogous expression of (9):

$$-\frac{1}{\beta}\log(\alpha)=(U^{int}(L=bound, R=on) - U^{int}(L=bound, R=off))$$
$$+(U^{int}(L=unbound, R=off) - U^{int}(L=unbound, R=on)) \quad (18)$$

As (18) indicates, the allosteric efficacy is a function the interaction energy between the states, and we have succeeded in expressing the thermodynamic allosteric efficacy as a function of local interactions in our simple two-component ligand/receptor system. However, this result is significantly more useful for considering multi-component systems if additional energetic symmetries are imposed by using an Ising model potential energy function. While these symmetries are not strictly realized in a biomolecular system, we will show that their application leads to concise analytic expressions that are qualitatively and quantitatively accurate as well for systems in which these symmetries are not present.

2.2. The Allosteric Ising Model (AIM) for Multicomponent Systems

The Ising model is a statistical mechanical model originally developed to describe phase behavior in ferromagnetic materials [21]. The Ising model, as well as Ising-like models, have since been applied to other complex systems with collective behavior [22,23], including cooperativity during folding [24–26] and in oligomeric assemblies [27,28]. In the Ising model, each particle has two states, corresponding to a spin state of up or down:

$$s_x = \begin{cases} -1 & X = \downarrow \\ 1 & X = \uparrow \end{cases} \quad (19)$$

The potential energy function of an n-component Ising model is:

$$U(X) = -\sum_{i=1}^{n} h_i s_i - \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{j_{ij}}{2} s_i s_j \quad (20)$$

In the Ising model, h_i is the potential energy of particle i due to the magnetic field, and j_{ij} is the spin coupling between particles i and j, where j_{ii} is taken to be 0. If the field term is taken to be site-specific, one can see that the field term can be considered to correspond to the conformational energy, and the spin coupling term to the pairwise interaction energy. We can rewrite the potential function as:

$$U(X) = \sum_{i=1}^{n} u_i^{conf} s_i + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{u_{i,j}^{int}}{2} s_i s_j \quad (21)$$

where u_i^{conf} is the conformational energy of component i, and $u_{i,j}^{int}$ is the interaction energy of components i and j. By using (21) for the potential energy function, we impose the following symmetries on the two-state components (with binary states represented by up and down arrows):

$$U^{conf}(X = \uparrow) = -U^{conf}(X = \downarrow)$$
$$U^{int}(X_i = \uparrow, X_j = \uparrow) = U^{int}(X_i = \downarrow, X_j = \downarrow) = -U^{int}(X_i = \uparrow, X_j = \downarrow) = -U^{int}(X_i = \downarrow, X_j = \uparrow) \quad (22)$$

For Ising models composed of several components and various interaction topologies, these symmetries allow for concise analytical expression for the allosteric efficacy and binding affinity. We will refer to these models as Allosteric Ising Models (AIMs).

Considering the analogy to the ligand(L)-receptor(R) systems and treating the *on/off* and *bound/unbound* states as *up/down* spins (see Figure 2A), the potential energy function according to (21) can be written as:

$$U(s_L, s_R) = u_L^{conf} s_L + u_R^{conf} s_R + u_{L,R}^{int} s_L s_R \quad (23)$$

As the interaction energy between the receptor and the ligand must be zero when the ligand is in the unbound state, we write an alternative non-Ising potential energy function where the interaction energy is 0 when the ligand is unbound:

$$U(s_L, s_R) = u_L^{conf} s_L + u_R^{conf} s_R + u_{L,R}^{int} \frac{s_L + 1}{2} s_R \quad (24)$$

This equation can be re-written as an Ising model potential energy function:

$$U(s_L, s_R) = u_L^{conf} s_L + \left(u_R^{conf} + \frac{u_{L,R}^{int}}{2} \right) s_R + \frac{u_{L,R}^{int}}{2} s_L s_R \quad (25)$$

Thus we will proceed with (23) despite the seemingly non-physical interaction, and later confirm that the relationships derived using this model accurately represent those of non-Ising systems. The allosteric efficacy using this potential energy function is:

$$\alpha \frac{p(L=\downarrow, R=\uparrow)}{p(L=\downarrow, R=\downarrow)} = \frac{p(L=\uparrow, R=\uparrow)}{p(L=\uparrow, R=\downarrow)} \quad (26)$$

and we can simplify (17) to:

$$\alpha = e^{-4\beta u_{L,R}^{int}} \quad (27)$$

Equation (27) indicates that in the Allosteric Ising Model for the ligand/receptor system ("ligand/receptor AIM"), the allosteric efficacy is simply a function of the ligand-receptor interaction energy term. Positive allostery (agonism) is attributed to negative interaction energy; negative allostery (inverse agonism) is attributed to positive interaction energy. Note that as the interaction energy between the ligand and receptor is related to the allosteric efficacy by a log transformation, we will use here the allosteric efficacy and interaction energy interchangeably, and specifically use interaction energy for visual representations, where the log scale is required.

The two-component model assumes that the protein is entirely rigid, with two global states. However, it is possible for the ligand to allosterically modulate multiple distinct allosteric sites (see Figure 2B). It is well known, for example, that GPCRs can signal through multiple downstream signaling pathways through coupling to various G protein subtypes and β-arrestin [29,30], and that different ligands can differentially activate these pathways [31,32]. Therefore it may be necessary to distinguish among multiple allosteric sites in the representation of receptor allostery. If we introduce two non-interacting allosteric sites, A_1 and A_2, we can write the potential energy function as:

$$U(L, A_1, A_2) = u_L^{conf} + u_{A_1}^{conf} + u_{A_2}^{conf} + u_{L,A_1}^{int} + u_{L,A_2}^{int} \quad (28)$$

Then the allosteric efficacy at a site is:

$$\alpha \frac{p\left(L=\downarrow, A_1=\uparrow\right)}{p\left(L=\downarrow, A_1=\downarrow\right)} = \frac{p\left(L=\uparrow, A_1=\uparrow\right)}{p\left(L=\uparrow, A_1=\downarrow\right)} \quad (29)$$

The probability of each state is the sum of the probability of two underlying states:

$$\alpha_{L,A_1} \frac{p\left(L=\downarrow, A_1=\uparrow, A_2=\uparrow\right) + p\left(L=\downarrow, A_1=\uparrow, A_2=\downarrow\right)}{p\left(L=\downarrow, A_1=\downarrow, A_2=\uparrow\right) + p\left(L=\downarrow, A_1=\downarrow, A_2=\downarrow\right)} = \frac{p\left(L=\uparrow, A_1=\uparrow, A_2=\uparrow\right) + p\left(L=\uparrow, A_1=\uparrow, A_2=\downarrow\right)}{p\left(L=\uparrow, A_1=\downarrow, A_2=\uparrow\right) + p\left(L=\uparrow, A_1=\downarrow, A_2=\downarrow\right)} \quad (30)$$

which is equal to:

$$\alpha_{L,A_1} \frac{e^{-\beta\left[-u_L^{conf}+u_{A_1}^{conf}+u_{A_2}^{conf}-u_{L,A_1}^{int}-u_{L,A_2}^{int}\right]} + e^{-\beta\left[-u_L^{conf}+u_{A_1}^{conf}-u_{A_2}^{conf}-u_{L,A_1}^{int}+u_{L,A_2}^{int}\right]}}{e^{-\beta\left[-u_L^{conf}-u_{A_1}^{conf}+u_{A_2}^{conf}+u_{L,A_1}^{int}-u_{L,A_2}^{int}\right]} + e^{-\beta\left[-u_L^{conf}-u_{A_1}^{conf}-u_{A_2}^{conf}+u_{L,A_1}^{int}+u_{L,A_2}^{int}\right]}}$$

$$= \frac{e^{-\beta\left[u_L^{conf}+u_{A_1}^{conf}+u_{A_2}^{conf}+u_{L,A_1}^{int}+u_{L,A_2}^{int}\right]} + e^{-\beta\left[u_L^{conf}+u_{A_1}^{conf}-u_{A_2}^{conf}+u_{L,A_1}^{int}-u_{L,A_2}^{int}\right]}}{e^{-\beta\left[u_L^{conf}-u_{A_1}^{conf}+u_{A_2}^{conf}+u_{L,A_1}^{int}-u_{L,A_2}^{int}\right]} + e^{-\beta\left[u_L^{conf}-u_{A_1}^{conf}-u_{A_2}^{conf}-u_{L,A_1}^{int}-u_{L,A_2}^{int}\right]}} \quad (31)$$

This reduces to:

$$\alpha_{L,A_1} = e^{-4\beta u_{L,A_1}^{int}} \quad (32)$$

which indicates that the allosteric efficacy of a ligand at an allosteric site is independent of other allosteric sites it modulates as well (provided the allosteric sites are not coupled through another interaction). In terms of receptor signaling, this analysis predicts that there could exist ligands with absolute bias for only one signaling pathway. This would require the downstream effectors (*e.g.,* the G proteins or β-arrestin for GPCRs) to interact with unique and independent allosteric sites.

2.3. Representation of allosteric propagation through specific regions within the protein

In addition to the existence of multiple allosteric sites, allosteric conformational coupling can be propagated through specific regions within the protein, often called "paths" or "channels". Using the AIM approach described here, we can expand the treatment of allostery to proteins with multiple structural components, where some components are allosterically regulated, and some others mediate the allosteric regulation. We begin with a three-component model, composed of the ligand L, a channel C, and an allosteric site A (see AIM represented in Figure 2C). The potential energy function is:

$$U\left(L, C, A\right) = u_L^{conf} s_L + u_C^{conf} s_C + u_A^{conf} s_A + u_{L,C}^{int} s_L s_C + u_{C,A}^{int} s_C s_A + u_{L,A}^{int} s_L s_A \quad (33)$$

The allosteric efficacy is then:

$$\alpha_{L,A_1} \frac{e^{-\beta\left[-u_L^{conf}+u_C^{conf}+u_A^{conf}-u_{L,C}^{int}-u_{L,A}^{int}+u_{C,A}^{int}\right]}+e^{-\beta\left[-u_L^{conf}+u_C^{conf}-u_A^{conf}-u_{L,C}^{int}+u_{L,A}^{int}-u_{C,A}^{int}\right]}}{e^{-\beta\left[-u_L^{conf}-u_C^{conf}+u_A^{conf}+u_{L,C}^{int}-u_{L,A}^{int}-u_{C,A}^{int}\right]}+e^{-\beta\left[-u_L^{conf}-u_C^{conf}-u_A^{conf}+u_{L,C}^{int}+u_{L,A}^{int}+u_{C,A}^{int}\right]}}$$

$$=\frac{e^{-\beta\left[u_L^{conf}+u_C^{conf}+u_A^{conf}+u_{L,C}^{int}+u_{L,A}^{int}+u_{C,A}^{int}\right]}+e^{-\beta\left[u_L^{conf}+u_C^{conf}-u_A^{conf}+u_{L,C}^{int}-u_{L,A}^{int}-u_{C,A}^{int}\right]}}{e^{-\beta\left[u_L^{conf}-u_C^{conf}+u_A^{conf}+u_{L,C}^{int}-u_{L,A}^{int}-u_{C,A}^{int}\right]}+e^{-\beta\left[u_L^{conf}-u_C^{conf}-u_A^{conf}-u_{L,C}^{int}-u_{L,A}^{int}+u_{C,A}^{int}\right]}}$$

(34)

Equation (34) simplifies to:

$$\alpha_{L,A}=e^{-4\beta u_{L,A}^{int}}\frac{\cosh\left(2\beta\left(u_{L,C}^{int}+u_{C,A}^{int}\right)\right)+\cosh\left(2\beta u_C^{conf}\right)}{\cosh\left(2\beta\left(u_{L,C}^{int}-u_{C,A}^{int}\right)\right)+\cosh\left(2\beta u_C^{conf}\right)} \quad (35)$$

where cosh is the hyperbolic cosine function:

$$\cosh(x)=\frac{e^x+e^{-x}}{2} \quad (36)$$

It should be noted that the exponential term in (35) is the *conditional allosteric efficacy*. The conditional allosteric efficacy can be written as the sum of weighted allosteric efficacies, with each allosteric efficacy conditioned on a different state of the channel and then weighted by the corresponding probability of that state:

$$\alpha_{L,A\big|C}=p\left(C=\uparrow\right)\alpha_{L,A\big|C=\uparrow}+p\left(C=\downarrow\right)\alpha_{L,A\big|C=\downarrow} \quad (37)$$

where for a given state, s, of C:

$$\alpha_{L,A\big|C=s}=\frac{p\left(L=\uparrow,A=\uparrow,C=s\right)p\left(L=\downarrow,A=\downarrow,C=s\right)}{p\left(L=\uparrow,A=\downarrow,C=s\right)p\left(L=\downarrow,A=\uparrow,C=s\right)} \quad (38)$$

Equation (38) simplifies to:

$$\alpha_{L,A\big|C}=e^{-4\beta u_{L,A}^{int}} \quad (39)$$

Comparing (39) with the allosteric efficacy of the two-component ligand/receptor system expressed in (27), it is clear that the conditional allosteric efficacies in the three-component system are simply the allosteric efficacies of the corresponding two-component systems.

We can then differentiate the allosteric efficacy contributed by the direct interaction of two components, *i.e.*, the conditional allosteric efficacy, from the indirect contributions, and write:

$$\alpha_{L,A} = \alpha_{L,A|C} \, \alpha_{L,A}^{indirect} \quad (40)$$

where the allosteric efficacy contributed by the indirect interaction is:

$$\alpha_{L,A}^{indirect,C} = \frac{\cosh\left(2\beta\left(u_{L,C}^{int} + u_{C,A}^{int}\right)\right) + \cosh\left(2\beta u_C^{conf}\right)}{\cosh\left(2\beta\left(u_{L,C}^{int} - u_{C,A}^{int}\right)\right) + \cosh\left(2\beta u_C^{conf}\right)} \quad (41)$$

Importantly, (41) provides a description of the allosteric efficacy as a function of the channel through which it is propagated. There are immediate inferences that can be drawn from this representation. First, the channel must have little preference for either one of its conformations, so that signaling through it can have a high intrinsic signal-to-noise ratio. Based on this inference, mutations that further stabilize the intrinsically preferred conformation of a channel will decrease the allosteric efficacy of a ligand, whereas mutations that destabilize that conformation will increase the allosteric efficacy. The existence of these two classes of mutations has immediate implications for the ability to test experimentally the role of specific domains in allosteric signaling. Second, because allosteric transmission through the channel depends on a balance between the channel's conformational energy and the interaction energy between the channel and ligand, and the channel and allosteric site, it follows that a low intrinsic signal-to-noise ratio can be overcome by an increased coupling of the ligand to the channel. Lastly, if the sign of the coupling of the ligand to the channel is opposite that of the channel to the allosteric site, the allosteric signal can be reversed. Consequently, a binding site on a protein that has been evolved for positive allostery by endogenous ligands, can be targeted as a site for negative allosteric modulation, and *vice versa*. It is well known that endogenous agonist-binding sites can be targeted by inverse-agonists, so this result is anchored in experimental evidence.

2.4. The Channel as a Chain of Interacting Structural Components

Comparison of (35) with (39) indicates that the allosteric efficacy can be written in terms of the conditional allosteric efficacies due to direct interactions:

$$\alpha_{L,A} = \alpha_{L,A|C} \frac{\cosh\left(\frac{1}{2}\log\left(\alpha_{L,C|A}\, \alpha_{C,A|L}\right)\right) + \cosh\left(2\beta u_C^{conf}\right)}{\cosh\left(\frac{1}{2}\log\left(\frac{\alpha_{L,C|A}}{\alpha_{C,A|L}}\right)\right) + \cosh\left(2\beta u_C^{conf}\right)} \quad (42)$$

In effect, the conditional allosteric efficacy is the signal-to-noise ratio for a single step in the signal propagation process, and the effective signal-to-noise ratio for the entire signal propagation system can be described by a non-linear function of all the constituent propagation steps.

Equation (42) can also be written as the effective interaction energy, $u_{L,A}^{int*}$:

$$u_{L,A}^{int*} = u_{L,A}^{int} - \frac{1}{4\beta}\log\left(\frac{\cosh\left(2\beta\left(u_{L,C}^{int} + u_{C,A}^{int}\right)\right) + \cosh\left(2\beta u_C^{conf}\right)}{\cosh\left(2\beta\left(u_{L,C}^{int} - u_{C,A}^{int}\right)\right) + \cosh\left(2\beta u_C^{conf}\right)}\right) \quad (43)$$

and thus as the sum of the direct and indirect interactions:

$$u_{L,A}^{int*} = u_{L,A}^{int} + u_{L,A}^{indirect,C} \quad (44)$$

It should be noted that the designation of channel versus allosteric site is purely an operational definition in which the site that performs the function of interest is referred to as the allosteric site. If both sites are functional, such as the example of two independent allosteric sites described above, and if they interact, we can rewrite (42) as:

$$\alpha_{L,A_1} = \alpha_{L,A_1|A_2} \frac{\cosh\left(\frac{1}{2}\log\left(\alpha_{L,A_2|A_1}\,\alpha_{A_1,A_2|L}\right)\right) + \cosh\left(2\beta u_{A_2}^{conf}\right)}{\cosh\left(\frac{1}{2}\log\left(\frac{\alpha_{L,A_2|A_1}}{\alpha_{A_1,A_2|L}}\right)\right) + \cosh\left(2\beta u_{A_2}^{conf}\right)} \quad (45)$$

The description of the allosteric efficacy as a function of the channel through which it is propagated, in (41), indicates that if the channel is a one-dimensional chain of interacting structural components, the allosteric efficacy is quickly diminished (it has been shown that the spin correlation function decays exponentially with distance in one-dimensional Ising models [21]). In Figure 3, the effective interaction energy between the first and last components of one-dimensional Ising chains with uniform conditional allosteric efficacies of 10, 100, 1,000, 10,000, and 100,000 are shown as a function of chain length. For weakly interacting systems, channels formed by structural components interacting in series do not appear to be good mediators of allosteric efficacy. The prevalence of multi-segment transmembrane signaling complexes may indicate an evolutionary mechanism to overcome the limitations of serial channels.

2.5. Comparison of Allosteric Propagation in Ising and Non-Ising Systems

As described in Section 2.2, the above analysis is made possible through the energetic symmetries imposed by the Ising model. However, it is unlikely these energetic symmetries exist in real allosteric proteins. Thus, it is important to consider how well the relationships derived from AIMs describe non-Ising two-state models, which are expected to be better representations of the types of interaction networks present in the biomolecular systems of interest.

To consider this problem, we sampled 100,000 non-Ising two-state allosteric systems with interaction energies and configurational energies sampled from normal distributions of mean 0 and standard deviation of β^{-1}, $3/\beta$, or $5/\beta$. The exact allosteric efficacies, calculated from the exact probabilities of each state, were then compared to the allosteric efficacies estimated from (42) using the direct allosteric efficacy terms. We should note that while

direct allosteric efficacies can be calculated for non-Ising model, the calculation of the configuration energy term followed:

$$2u_C^{conf} \approx U^{conf}(C=\uparrow) - U^{conf}(C=\downarrow) \quad (46)$$

As above, we addressed problems that may arise from the non-physical interaction energy between unbound ligand and the protein by setting to 0 all interaction energies with the unbound ligand. Results of these calculations are shown in Figure 4, where the corresponding effective interaction energies have been used for clarity. Our calculations indicate that (42) is a good estimate of the true allosteric efficacy in non-Ising systems in which the allosteric efficacy is high (see Figure 4A). As the standard deviation on the energy term distribution increases, and more systems have significant deviation from Ising-like behavior, two distinct groups of false positives (exact effective interaction energy is 0 but estimated interaction energy is non-zero) and true negatives (exact effective interaction energy is non-zero but estimated interaction energy is 0) do appear, but the sign of the allosteric modulation is conserved (see Figures 4B,C).

That the model maintains high accuracy for systems with high allosteric efficacy in spite of the two groups of inaccuracy (*i.e.*, false positives and true negatives), suggests that this model should reflect many of the qualitative and quantitative properties of actual allosteric systems.

2.5. A Relation of AIMs to the Structural Dynamics Analysis of Biomolecular Function

Efforts to identify allosteric sites and channels in the structures of functional biomolecules have utilized estimates of correlation or mutual information between the structural dynamics of known allosteric sites and candidate modulation sites or channels, most often based on the analysis of molecular dynamics (MD) trajectories [33,34,18,19] or elastic network models (ENMs) [35,36]. Equation (43) indicates that structural components that can act as channels will have high effective interaction energy with known allosteric sites (e.g., $u_{C,A}^{int}$), and the Shannon-Hartley theorem, (8), suggests that the allosteric efficacy can be related to the mutual information via the channel capacity. It is not clear, however, how this relates to the mutual information that is evaluated from an MD simulation. As we and others have used mutual information successfully to interpret the structural dynamics and allostery from MD trajectories [18–20], it is interesting to test the use of mutual information as an identifier of allostery in the context of AIMs. To this end we calculated the *symmetric uncertainty* [37], a normalized variant of the mutual information, between each component in two-component Ising models and two-component non-Ising models, and compared the *symmetric uncertainty* to the absolute interaction energy. The *symmetric uncertainty* (SU) between components is:

$$SU(X_i, X_j) = \frac{2I(X_i, X_j)}{H(X_i) + H(X_j)} \quad (47)$$

where I is the mutual information:

$$I\left(X_i, X_j\right) = H\left(X_i\right) + H\left(X_j\right) - H\left(X_i, X_j\right) \quad (48)$$

and H is the Shannon entropy:

$$H\left(X\right) = -\sum p\left(X\right) \log\left(p\left(X\right)\right) \quad (49)$$

We generated 100,000 two-component Ising systems and 100,000 two-component non-Ising systems with energy terms sampled from a normal distribution with mean 0 and standard deviation of 1, and calculated the symmetric uncertainty and allosteric efficacy of each. We find that the symmetric uncertainty enforces a lower limit on the allosteric efficacy, and allosteric efficacy increases with higher symmetric uncertainty (see Figure 5). Thus, mutual information is a good predictor of allosteric activity in the two-state models explored here. The use of mutual information in systems that are not two-state will be discussed further below.

2.6. AIMs and Multiple Allosteric Channels

Many proteins have been suggested to have multiple allosteric channels [38]. Assuming that the channels are independent, careful algebra (not shown) reveals that to study the allosteric efficacy of a multi-channel system one can iteratively replace the direct interaction energy term with a direct interaction and indirect interaction of the same effective interaction energy. The effective interaction energy due to multiple independent channels is additive:

$$u_{L,A}^{int*} = u_{L,A}^{int} + \sum_{i=1}^{N} u_{L,A}^{indirect,C_N} \quad (50)$$

and the allosteric efficacy is then multiplicative:

$$\alpha_{L,A} = \alpha_{L,A \mid \left\{C_1, \ldots, C_N\right\}} \prod_{i=1}^{N} \alpha_{L,A}^{indirect,C_N} \quad (51)$$

This formally obvious result reveals the advantage of multiple channels in an allosteric protein: perturbations such as mutations that disrupt the conformational stability of one channel will not abolish allosteric function completely. Many parallel weak channels introduce significant robustness when compared to the allosterically equivalent single strong channel built in series, because the latter is completely eliminated by disruption of even a single interaction between two of its structural components.

To test the ability of Equation (51) to reflect accurately the behavior of non-Ising systems, we again constructed 100,000 two- and three-channel non-Ising allosteric systems using the methodology described for single channel systems, and compared the resulting allosteric efficacy to that calculated using (51) (see Figure 6). Again, we find good agreement between the estimates using (51) and the exact calculated efficacies, although the accuracy is slightly reduced as the number of channels increases from two to three.

Because it is unlikely that allosteric proteins consist of absolutely independent channels, we explored the effect of interaction between channels through the use of two AIMs: one two-channel system where both channels provide equal magnitude positive allosteric coupling, and one two-channel system where both channels are of equal magnitude but opposite direction. The allosteric efficacy was calculated for each system as a function of the interaction energy between the two channels of allostery for ligands that are coupled to one, or both channels.

As depicted in Figure 7, we found that when two channels mediating positive allosteric modulation have a negative interaction energy, the allosteric efficacy of the ligand is increased, even if the ligand only interacts with one channel (Figure 7A). This is not unexpected; the second channel acts as an indirect channel from the first channel to the allosteric site and additionally multiplies the allosteric efficacy of the channel. However, if the ligand interacts with both channels, the allosteric efficacy is not the square of the allosteric efficacy of binding to one channel as would be for two identical, independent channels. This is because the interaction of the ligand with the first channel has already partially shifted the conformational distribution of the second channel, decreasing its channel efficacy by effectively increasing its intrinsic conformational preference (and thus its intrinsic signal-to-noise).

For the second two-channel system, with channels providing allosteric coupling in opposite directions, we find that when the interaction energy between the channels is negative, there is decreased allosteric efficacy for the ligand in either channel, whereas positive interaction energy between the channels leads to increased allosteric efficacy (Figure 7B). From the perspective of the positive channel, if the channels have a negative interaction energy, the second (negative) channel is an indirect channel that flips the sign of the allosteric signal; this leads to reduced overall allosteric efficacy due to negation. However, if the two channels have a positive interaction energy, the signal through the second channel is flipped twice and left unchanged, leading to increased allosteric efficacy. Interestingly, if the ligand interacts with both channels equally, the effective interaction energy from this pair of channels is 0, independent of the interactions between the channels. In a receptor with these characteristics, antagonists could interact with each channel without conformational preference for the channel, or interact with both channels with the same sign, leading to no allosteric signal.

2.7. Illustration of AIM-Based Analysis of Allosteric Coupling Mechanisms: The Asymmetric D2 Receptor Homodimeric Signaling Complex

The application of the new formalism based on AIMs was used thus far to represent small, ideal systems in order to extract insights into the physics of allostery on a conceptual level. To examine the practical implementation of AIMs for real allosteric proteins of biological interest, we chose to construct AIMs consisting of a small number of structural components where the numerical calculations of allosteric properties can be performed easily. Such use of AIMs as a coarse-grain level of representation is advantageous in testing hypotheses about the underlying structural mechanisms of real allosteric proteins. This concept is illustrated here with the example of a well-characterized GPCR dimer system.

We constructed a model of asymmetric signaling in the dopamine D2 receptor (D2R) homodimer, based on the structural model of the asymmetric dimer and the constructs used to explore its function that were published recently [39]. Because D2R can signal as both a monomer and a homodimer, a novel experimental construct developed in the Javitch lab [39] was required to make possible the characterization of the dimer as a signaling unit. The results demonstrated experimentally that the D2R homodimer cannot signal through each monomer simultaneously, but instead signals through a single protomer at a time in an asymmetric manner (the signaling protomer will be referred to as "protomer A"). Furthermore, the results indicate that the function of the protomers is characterized by negative cooperativity: the stabilization of the *on* state of the non-signaling monomer ("protomer B") by agonist biding decreases signaling by protomer A, whereas the stabilization of the *off* state of protomer B by the binding of an inverse agonist increases signaling by protomer A. Lastly, it is shown in [39] that perturbations known to completely disrupt activation in the monomer, including: (i)-ablation of ligand binding, (ii)-removal of intracellular loop 3 (IL3), and (iii)-mutations introduced in (a)-intracellular loop 2 (IL2), (b)-the conserved DRY motif, and (c)-the conserved NPxxY motif, all disrupt activation in the homodimer when applied to protomer A. Unexpectedly, however, the perturbations in (iii) also disrupt activation when applied to protomer B.

A molecular model of the homodimer complex with the G protein that senses the activation of the receptor was constructed in [39] to explain the experimental results in a structural context. The template for this model was the active state crystal structure of another GPCR, rhodopsin, bound to its G protein, transducin. In this molecular model the interface of the homodimer involves the 4th transmembrane segment (TM4), and the G protein interacts with the signaling protomer A through IL3, IL2, and helix 8 (H8), while protomer B interacts through its IL2 and H8 (see Figure 8). We used AIMs as described below to explore the feasibility of the allosteric properties proposed for this structural model.

Based on the experimental measurements of activation, an AIM representing the homodimer was constructed starting with a model for a signaling monomer (monomer A) and a G protein that can bind this monomer and become activated. Since the IL2, DRY, and NPxxY mutations behave identically in the experiments, we represented all three as a single structural component termed *the conserved binding motifs* (CBMs), due to their role in G protein activation by the GPCR [40–44]. In this AIM (see Figure 8A), the signaling monomer is composed of the following structural components: a ligand that can bind and unbind, a transmembrane domain, and two intracellular regions (IL3 and the CBMs); the G protein is composed of a structural component that can bind and unbind the signaling monomer, and one that can be activated. The conformational energies of the components of each protomer were chosen to prefer the *off* state ($u^{conf} = 1$), and the interaction energies between all components were negative such that they preferred to be in the same state ($u^{int} = -1$). We find that this coarse grained model responds as expected to agonists, antagonists, and inverse agonists (see Figure 8B). To create a homodimer with negative cooperativity, we then added to the AIM a negative cooperativity between the one monomer that can bind G protein (which is now protomer A) and one that cannot (protomer B), represented as a positive interaction energy between their transmembrane domains (see Figure 8C). We then

calculated the allosteric efficacy for the homodimer when protomer A was bound to agonist and protomer B was simultaneously bound to either an agonist, an antagonist, or an inverse agonist. This model reproduces the observed negative cooperativity (See Figure 8D)

To explore the effects of removing IL3 and introducing the CBM mutations, we constructed AIMs with the perturbations modeled as either: i) stabilizing the *off* state of the mutated structural component, ii) stabilizing its *on* state, or iii) reducing the interaction energy between the structural component and the G protein to 0. Modeling the two perturbations in protomer A by imposing (i) or (iii), reduced activation as expected. However, stabilizing the *off* state of IL3 in protomer B increases activation in our model when it should have no effect, indicating that treating the IL3 mutation such that it eliminates interaction between IL3 and the G protein is a better model. On the other hand, treating the CBM perturbation in protomer B as stabilizing the *off* state leads to more activation, so that the effect of the mutation cannot be explained without an interaction between the CBM in protomer B and the G protein. To reconcile these effects in the model, we assumed that protomer B and the G protein bind in a state-independent way (the G protein's state independent binding is represented by $u_{G_{binding}}^{conf}$ in the AIM), and modeled the CBM mutation effect as further decreasing state-independent binding. We find that if $u_{G_{binding}}^{conf}$ is increased from 1 to 2, allosteric efficacy is reduced (see Figure 8D). The finding that state-independent interactions between the G protein and CBMs on both protomer A and protomer B are required for activation is in full agreement with the structural model of the dimer as presented [39], in which not only protomer A, but also IL2 and H8 from protomer B interact with the G protein directly. As this structural information was not used in the construction of the AIMs, the prediction from the allosteric model underscores the ability of the AIMs-based approach in this illustration to connect the representation of allostery with the structural context of the modeled biomolecular systems.

3. Conclusions

We have explored models of biomolecular allostery through the use of Allosteric Ising Models (AIMs) in order to develop a quantitative theoretical description that bridges the features of the structural components and their interactions, to the thermodynamic allosteric parameters. From this perspective, we show first that the allosteric efficacy is the steady state signal-to-noise ratio for the ligand signal through the corresponding noisy receptor. We find that the allosteric efficacy, or the corresponding effective interaction energy, between two allosterically coupled sites can be expressed in terms of the conformational and interaction energies of the constituent parts for many small systems and interaction motifs. This formulation allows us to show that the allosteric efficacy is the product of the indirect allosteric efficacies through independent pathways, suggesting a mechanism by which biomolecular systems have evolved to be robust to mutation. While the equations were derived here using the Ising model to make use of symmetries in the potential energy function, we show that the model can produce good estimates of the allosteric properties of non-Ising two-state pairwise interaction models as well.

A general inference from the use of AIMs as discussed here is that the results can suggest some constraints on the design principles of allosteric proteins. Thus, we find that it is more efficient and more robust to use multiple parallel channels that are individually weak than to use a single series channel that is strong, and that interactions between the parallel channels can additionally increase allosteric efficacy. From a structural perspective it is possible to surmise that α-helices behave as strong serial channels, where as β-sheets behave more like coupled parallel channels that are individually weak. Indeed, it has been shown that significant long-distance correlations exist in β-sheets [45], but little work has been done to study the connection of the properties of these fundamental units of protein structure to their involvement in known allosteric mechanisms. Understanding the allosteric properties of such structural components and common structural motifs from the perspective shown here offers valuable insight into how the wide array of allosteric proteins observed in nature could have been obtained from the limited number of amino acids and folding motifs.

The illustration of the application of AIMs to the D2 receptor homodimer was successful in producing an allosteric model that predicted structural details of molecular interactions. However, it is important to note that the AIM framework assumes that structural components within biomolecular systems exhibit two-state behavior. While this assumption has been used widely in the study of GPCRs and transporters (e.g., the proposed "rocking bundle" mechanism [46,47]), experimental and computational studies indicate [30,48–54] that the character of the conformational space sampled by these molecular machines is not strictly two-state as is often assumed. The principles demonstrated in this manuscript are not mathematically transferable directly to models where structural components require representation by: i)-more than two discrete states, or ii)-continuous states in one or more dimensions. The study of the more complex systems necessitates a more general approach such as the N-body information theoretical analysis we have previously developed [18,55]. We have used such an N-body Information Theory (NbIT) analysis to identify allosteric channels and collective behavior in both transporters [18] and GPCRs [55]. To address the more complex properties of large allosteric systems such as the complex biomolecules responsible for cell function, it may be necessary to formulate a generalization of the NbIT model that allows arbitrary allosteric systems to be constructed and explored in the manner in which the AIMs were analyzed here.

Acknowledgments

References

1. Shi Y. A Glimpse of Structural Biology through X-Ray Crystallography. Cell. 2014; 4:995–1014. [PubMed: 25416941]

2. Markwick PRL, Malliavin T, Nilges M. Structural biology by NMR: Structure, dynamics, and interactions. PLoS Comput Biol. 2008; 4:e1000168. [PubMed: 18818721]

3. Bai X, McMullan G, Scheres SH. How cryo-EM is revolutionizing structural biology. Trends Biochem Sci. 2015; 40:49–57. [PubMed: 25544475]

4. Ha T. Single-molecule fluorescence resonance energy transfer. Methods. 2001; 25:78–86. [PubMed: 11558999]

5. Sahu ID, McCarrick RM, Lorigan GA. Use of electron paramagnetic resonance to solve biochemical problems. Biochemistry. 2013; 52:5967–5984. [PubMed: 23961941]

6. Adcock SA, McCammon JA. Molecular dynamics: Survey of methods for simulating the activity of proteins. Chem Rev. 2006; 106:1589–1615. [PubMed: 16683746]

7. Bahar I, Lezon TR, Yang LW, Eyal E. Global dynamics of proteins: bridging between structure and function. Annu Rev Biophys. 2010; 39:23–42. [PubMed: 20192781]

8. Gether U. Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. Endocr Rev. 2000; 21:90–113. [PubMed: 10696571]

9. Monod J, Changeux JP, Jacob F. Allosteric proteins and cellular control systems. J Mol Biol. 1963; 6:306–329. [PubMed: 13936070]

10. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? Proteins. 2004; 57:433–43. [PubMed: 15382234]

11. Makhlynets OV, Raymond EA, Korendovych IV. Biochemistry. 2015; 54:1444–56. [PubMed: 25642601]

12. Hilser VJ, Wrabl JO, Motlagh HN. Structural and energetic basis of allostery. Annu Rev Biophys. 2012; 41:585–609. [PubMed: 22577828]

13. Tsai C, Nussinov R. A Unified View of "How Allostery Works.". PLoS Comput Biol. 2014; 10:e1003394. [PubMed: 24516370]

14. Leff P. The two-state model of receptor activation. Trends Pharmacol Sci. 1995; 16:89–97. [PubMed: 7540781]

15. Fenton AW. Allostery: an illustrated definition for the "second secret of life.". Trends Biochem Sci. 2008; 33:420–425. [PubMed: 18706817]

16. Shannon CE. A Mathematical Theory of Communication. Bell Syst Tech J. 1948; 27:623–656.

17. Shannon CE. Communication in the presence of noise. Proc Inst Radio Eng. 1949; 37:10–21.

18. LeVine MV, Weinstein H. NbIT—A New Information Theory-Based Analysis of Allosteric Mechanisms Reveals Residues that Underlie Function in the Leucine Transporter LeuT. PLoS Comput Biol. 2014; 10:e1003603. [PubMed: 24785005]

19. Bowman GR, Geissler PL. Equilibrium fluctuations of a single folded protein reveal a multitude of potential cryptic allosteric sites. Proc Natl Acad Sci USA. 2012; 109:11681–11686. [PubMed: 22753506]

20. Gasper P, Fuglestad B. Allosteric networks in thrombin distinguish procoagulant vs. anticoagulant activities. Proc Natl Acad Sci USA. 2012; 109:21216–21222. [PubMed: 23197839]

21. Ising E. Beitrag zur theorie des ferromagnetismus. Zeit Phys A Hadron Nucl. 1925; 31:253–258.

22. Hopfield J. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci USA. 1982; 79:2554–2558. [PubMed: 6953413]

23. Machta BB, Papanikolaou S, Sethna JP, Veatch SL. Minimal model of plasma membrane heterogeneity requires coupling cortical actin to criticality. Biophys J. 2011; 100:1668–1677. [PubMed: 21463580]

24. Muñoz V, Thompson PA, Hofrichter J, Eaton WA. Folding dynamics and mechanism of beta-hairpin formation. Nature. 1997; 390:196–199. [PubMed: 9367160]

25. Vorov OK, Livesay DR, Jacobs DJ. Helix/coil nucleation: A local response to global demands. Biophys J. 2009; 97:3000–3009. [PubMed: 19948130]

26. Vorov OK, Livesay DR, Jacobs DJ. Nonadditivity in conformational entropy upon molecular rigidification reveals a universal mechanism affecting folding cooperativity. Biophys J. 2011; 100:1129–1138. [PubMed: 21320459]

27. Bray D, Duke T. Conformational spread: the propagation of allosteric states in large multiprotein complexes. Annu Rev Biophys Biomol Struct. 2004; 33:53–73. [PubMed: 15139804]

28. Graham I, Duke T. Dynamic hysteresis in a one-dimensional Ising model: Application to allosteric proteins. Phys Rev E. 2005; 71:061923.

29. Perez DM, Karnik SS. Multiple signaling states of G-protein-coupled receptors. Pharmacol Rev. 2005; 57:147–161. [PubMed: 15914464]

30. Kahsai AW, Xiao K, Rajagopal S, Ahn S, Shukla AK, Sun J, Oas TG, Lefkowitz RJ. Multiple ligand-specific conformations of the β2-adrenergic receptor. Nat Chem Biol. 2011; 7:692–700. [PubMed: 21857662]

31. Urban JD, Clarke WP, Von Zastrow M, Nichols DE, Kobilka B, Weinstein H, Javitch JA, Roth BL, Christopoulos A, Sexton PM, Miller KJ, Spedding M, Mailman RB. Functional Selectivity and Classical Concepts of Quantitative Pharmacology. J Pharmacol Exp Ther. 2007; 320:1–13. [PubMed: 16803859]

32. Kenakin T. Functional selectivity and biased receptor signaling. J Pharmacol Exp Ther. 2011; 336:296–302. [PubMed: 21030484]

33. Sethi A, Eargle J, Black AA, Luthey-Schulten Z. Dynamical networks in tRNA: Protein complexes. Proc Natl Acad Sci USA. 2009; 106:6620–6625. [PubMed: 19351898]

34. Pandini A, Fornili A, Fraternali F, Kleinjung J. Detection of allosteric signal transmission by information-theoretic analysis of protein dynamics. FASEB J. 2012; 26:868–881. [PubMed: 22071506]

35. Ming D, Wall ME. Allostery in a coarse-grained model of protein dynamics. Phys Rev Lett. 2005; 95:1–4.

36. Su JG, Qi LS, Li CH, Zhu YY, Du HJ, Hou YX, Hao R, Wang JH. Prediction of allosteric sites on protein surfaces with an elastic-network-model-based thermodynamic method. Phys Rev E. 2014; 90:1–10.

37. Witten, IH.; Eibe, F.; Hall, MA. Data Mining Practical Machine Learning Tools and Techniques. Morgan Kaufmann; San Francisco, CA, USA: 2005. p. 664

38. Del Sol A, Tsai CJ, Ma B, Nussinov R. The origin of allosteric functional modulation: multiple pre-existing pathways. Structure. 2009:17, 1042–1050.

39. Han Y, Moreira IS, Urizar E, Weinstein H, Javitch JA. Allosteric communication between protomers of dopamine class A GPCR dimers modulates activation. Nat Chem Biol. 2009; 5:688–695. [PubMed: 19648932]

40. Moro O, Lameh J, Högger P, Sadée W. Hydrophobic amino acid in the i2 loop plays a key role in receptor-G protein coupling. J Biol Chem. 1993; 268:22273–22276. [PubMed: 8226735]

41. Ballesteros JA, Jensen AD, Liapakis G, Rasmussen SGF, Shi L, Gether U, Javitch JA. Activation of the β2-Adrenergic Receptor Involves Disruption of an Ionic Lock between the Cytoplasmic Ends of Transmembrane Segments 3 and 6. J Biol Chem. 2001; 276:29171–29177. [PubMed: 11375997]

42. Fritze O, Filipek S, Kuksa V, Palczewski K, Hofmann KP, Ernst OP. Role of the conserved NPxxY(x)5,6F motif in the rhodopsin ground state and during activation. Proc Natl Acad Sci USA. 2003; 100:2290–2295. [PubMed: 12601165]

43. Shan J, Khelashvili G, Mondal S, Mehler EL, Weinstein H. Ligand-Dependent Conformations and Dynamics of the Serotonin 5-HT 2A Receptor Determine Its Activation and Membrane-Driven Oligomerization Properties. PLoS Comput Biol. 2012; 8:e1002473. [PubMed: 22532793]

44. Han DS, Wang SX, Weinstein H. Active state-like conformational elements in the beta2-AR and a photoactivated intermediate of rhodopsin identified by dynamic properties of GPCRs. Biochemistry. 2008; 47:7317–7321. [PubMed: 18558776]

45. Fenwick RB, Orellana L, Esteban-Martín S, Orozco M, Salvatella X. Correlated motions are a fundamental property of β-sheets. Nat Commun. 2014; 5:4070. [PubMed: 24915882]

46. Forrest LR, Rudnick G. The rocking bundle: a mechanism for ion-coupled solute flux by symmetrical transporters. Physiology (Bethesda). 2009; 24:377–386. [PubMed: 19996368]

47. Forrest LR, Zhang YW, Jacobs MT, Gesmonde J, Xie L, Honig BH, Rudnick G. Mechanism for alternating access in neurotransmitter transporters. Proc Natl Acad Sci USA. 2008; 105:10338–10343. [PubMed: 18647834]

48. Nygaard R, Zou Y, Dror RO, Mildorf TJ, Arlow DH, Manglik A, Pan AC, Liu CW, Fung JJ, Bokoch MP, Thian FS, Kobilka TS, Shaw DE, Mueller L, Prosser RS, Kobilka BK. The dynamic process of β(2)-adrenergic receptor activation. Cell. 2013; 152:532–542. [PubMed: 23374348]

49. Provasi D, Artacho MC, Negri A, Mobarec JC, Filizola M. Ligand-induced modulation of the free-energy landscape of G protein-coupled receptors explored by adaptive biasing techniques. PLoS Comput Biol. 2011; 7:e1002193. [PubMed: 22022248]

50. Kazmier K, Sharma S, Quick M, Islam SM, Roux B, Weinstein H, Javitch JA, McHaourab HS. Conformational dynamics of ligand-dependent alternating access in LeuT. Nat Struct Mol Biol. 2014; 21:472–479. [PubMed: 24747939]

51. Shi L, Quick M, Zhao Y, Weinstein H, Javitch JA. The mechanism of a neurotransmitter:sodium symporter--inward release of Na$^+$ and substrate is triggered by substrate in a second binding site. Mol Cell. 2008; 30:667–677. [PubMed: 18570870]

52. Zhao C, Stolzenberg S, Gracia L, Weinstein H, Noskov S, Shi L. Ion-controlled conformational dynamics in the outward-open transition from an occluded state of LeuT. Biophys J. 2012; 103:878–888. [PubMed: 23009837]

53. Zhao Y, Terry D, Shi L, Weinstein H, Blanchard SC, Javitch JA. Single-molecule dynamics of gating in a neurotransmitter transporter homologue. Nature. 2010; 465:188–193. [PubMed: 20463731]

54. Zhao Y, Terry DS, Shi L, Quick M, Weinstein H, Blanchard SC, Javitch JA. Substrate-modulated gating dynamics in a Na+-coupled neurotransmitter transporter homologue. Nature. 2011; 474:109–113. [PubMed: 21516104]

55. LeVine, M.; Perez-Aguilar, J.; Weinstein, H. N-body Information Theory (NbIT) Analysis of Rigid-Body Dynamics in Intracellular Loop 2 of the 5-HT2A Receptor. Proceedings of of International-Work Conference on Bioinformatics and Biomedical Engineering 2014 (IWWBIO 2014); Granada, Spain. 7–9 April 2014; p. 1190-1201.
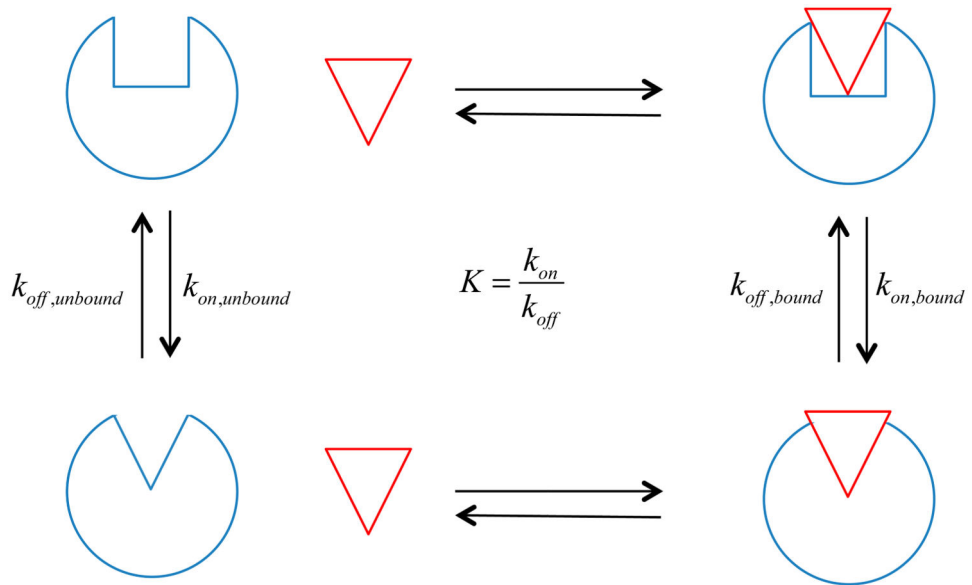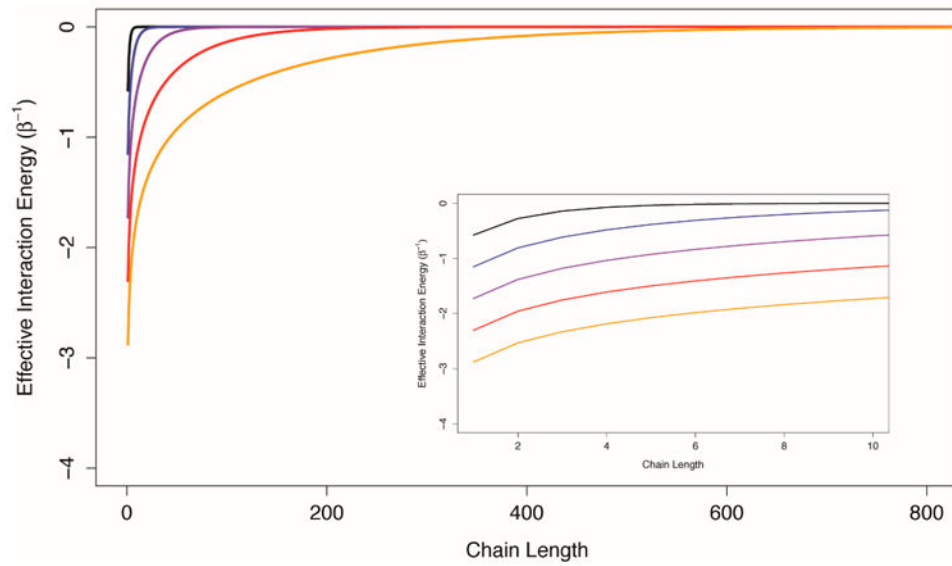
Figure 1.
Thermodynamic cycle of a two-state ligand/receptor activation reaction. The receptor (blue circle) has an *on* and an *off* state (square and triangle indentations, respectively), both of which can bind a ligand (red triangle). The kinetic parameters are shown for the two equilibria of interest.
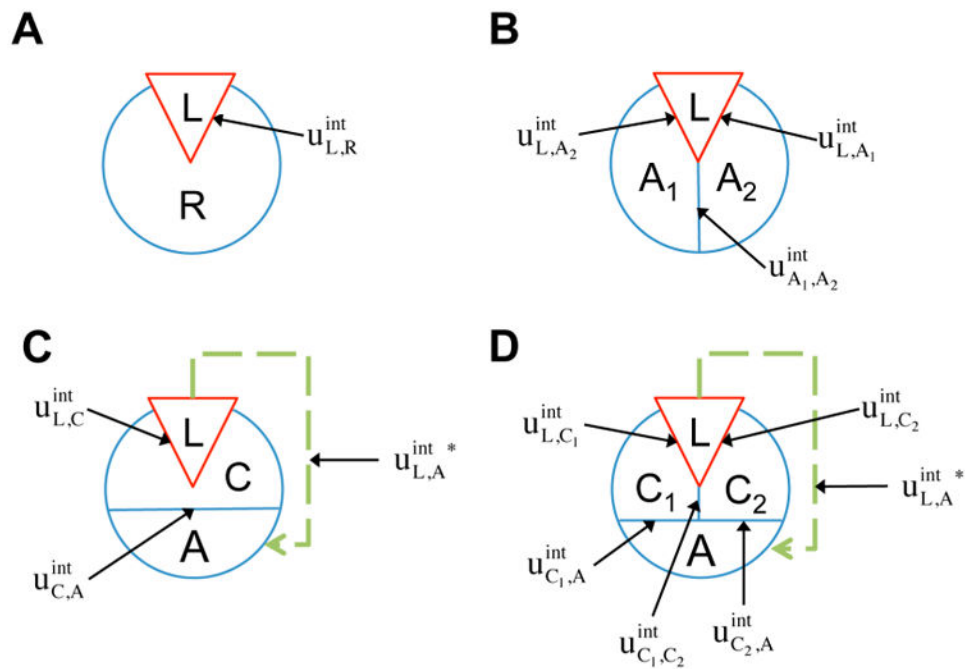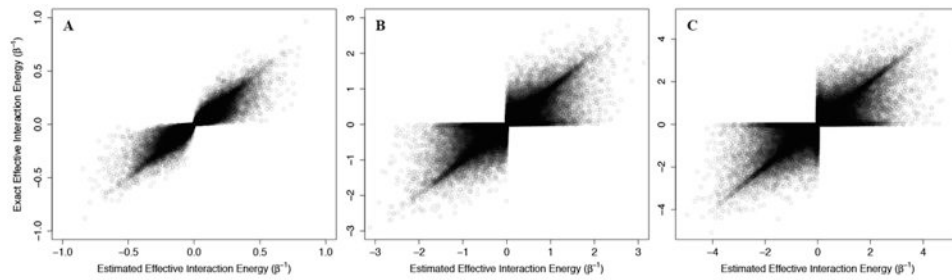
Figure 2.
Schematic representations of Allosteric Ising models (AIMs). In the four AIMs analyzed here the ligand, L, is represented as a red triangle, and the protein is the blue circle subdivided into various constituent structural components. Lines separating ligand from protein or protein structural components from each other are labeled with the appropriate interaction energy term (as used in the text). Allosteric effective interactions are represented with green dotted lines. (**A**): The simple two-component ligand/receptor system. (**B**): A three-component ligand/receptor system with two allosteric sites, A_1 and A_2. (**C**): A three-component ligand/receptor system with one channel, C, coupling the ligand and the allosteric site A. (**D**): A four-component ligand/receptor system with two channels, C1 and C2, coupling the ligand and the allosteric site A.
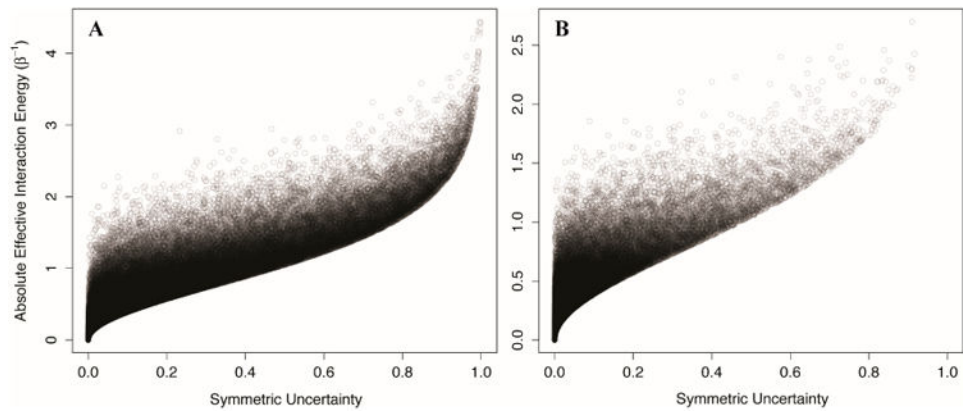
Figure 3.
The effective interaction energy through serial channels. Effective interaction energies of the first and last components of one-dimensional Ising chains are plotted as a function of chain length for conditional allosteric efficacy values of 10 (black), 100 (blue), 1000 (purple) 10,000 (red) and 100,000 (orange). The inset shows detail for short chain lengths. The effective interaction energy is seen to decay exponentially with channel length.
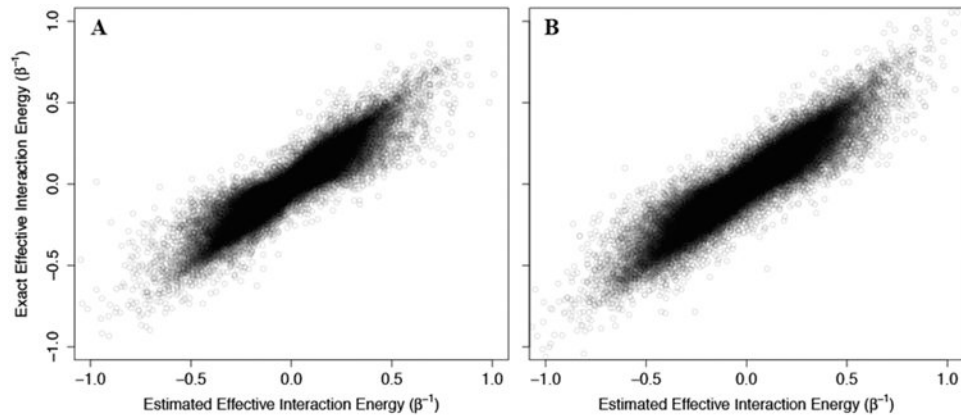
Figure 4.
Using the Ising model to estimate effective interaction energies in non-Ising three-component/two-state systems. The exact effective interaction energies of 100,000 three-component/two-state non-Ising systems are plotted against the effective interaction energy estimated using the equations derived for the three-component Ising model (see (42)). The systems are generated using energy terms sampled from a normal distribution of mean 0 and standard deviation of $1/\beta$ (**A**), $3/\beta$ (**B**), and $5/\beta$ (**C**) and the points are plotted with 10% opacity.
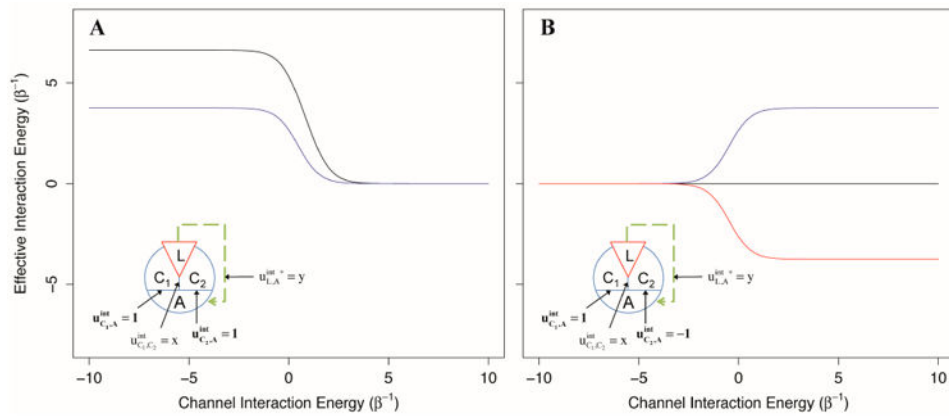
Figure 5.
Calculated *mutual information* between the channel and allosteric sites sets a lower bound on the allosteric efficacy. The symmetric uncertainty between the two components is plotted against the absolute effective interaction energy for 100,000 two-component/two-state non-Ising models (**A**), and two-component Ising models (**B**). The systems are generated using energy terms sampled from a normal distribution of mean 0 and standard deviation of $1/\beta$, and the points are plotted with 10% opacity.
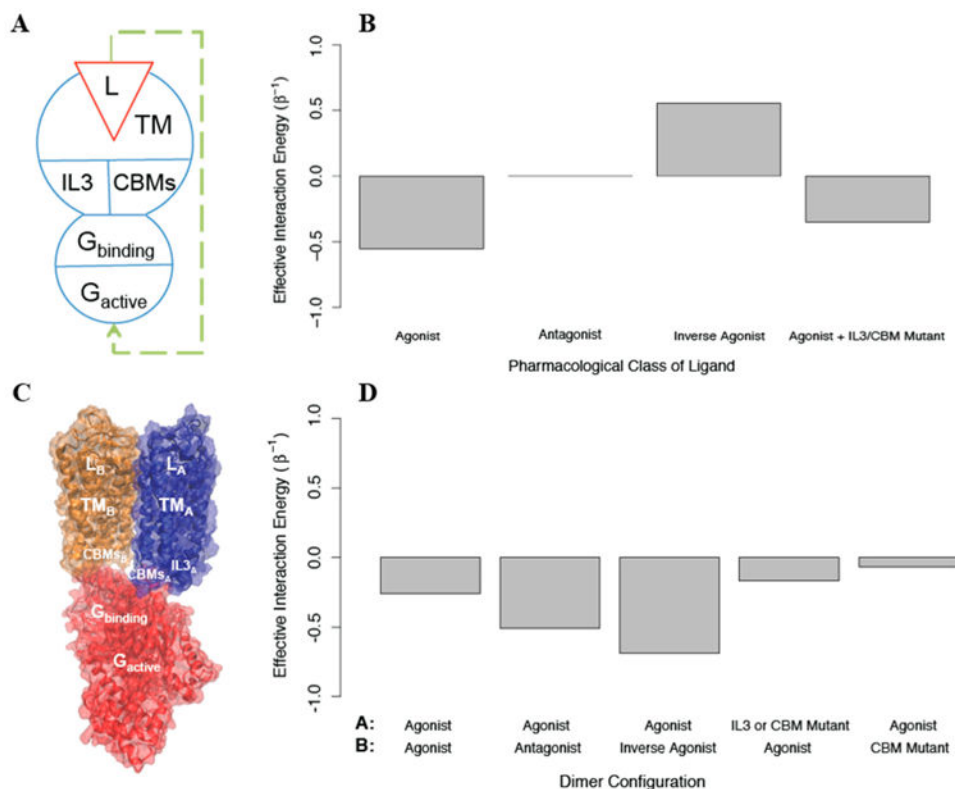
Figure 6.
Relation of effective interaction energies in non-Ising two-state systems with multiple independent channels to estimates from the corresponding Ising model. The exact effective interaction energies of 100,000 two-state non-Ising system is plotted against the effective interaction energy estimated using the equations derived for the n-channel Ising model (Equation (51)) for two (**A**), and three (**B**) independent channels. The systems are generated using energy terms sampled from a normal distribution of mean 0 and standard deviation of $1/\beta$, and the points are plotted with 10% opacity.

Figure 7.
The effective interaction energy of a two-channel AIM as a function of the interaction energy between the channels. (**A**): The two-channel system in which each channel contributes to positive allosteric modulation is shown for a ligand that interacts with one channel (blue) or both channels (black). (**B**): A two-channel system with one positive allosteric channel and one negative allosteric channel is shown for a ligand that interacts only with the positive channel (blue), only with the negative channel (red), or both channels (black). The effect of interactions between channels is seen to modify significantly the allosteric signal transduction.

Figure 8.
Analysis of the AIM for a well-characterized asymmetric D2 homodimer of the dopamine D2 receptor (D2R). (**A**): The D2R monomer AIM. (**B**): The effective interaction energy calculated for the D2R monomer AIM is presented for ligands that are agonists, antagonists, and inverse agonists, and also for the mutation of either IL3 or the *conserved binding motifs* (CBMs). (**C**): A molecular model of the homodimer obtained as described in the text, is shown with each AIM domain labeled in white on the structural representation. Protomer A is in blue, protomer B is in orange, and the G protein is in red. (**D**): The effective interaction energy for the D2R homodimer AIM is presented for different combinations of the states of protomer A (indicated by **A** in the top row) and those of protomer B in the dimer (**B**, bottom row).