Making sense of chaos: uncovering the mechanisms of conformational entropy

Stephanie A. Wankowicz[1], James S. Fraser[1]

1) Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA 94158, USA.

Abstract

By converting a disordered polymer into a globular structure, protein folding reduces many conformational degrees of freedom, resulting in a significant conformational entropy penalty. Nonetheless, residual entropy persists in the protein's native state as it fluctuates between thermally accessible conformations. Here, we review biophysical evidence, primarily from NMR studies, for how conformational entropy modulates the free energy of ligand binding and catalysis. The major theme that emerges is that selection based on free energy has converged on mechanisms to mitigate the effects of entropy loss during crucial functions like binding or catalysis. The modulation of conformational entropy occurs primarily via two main mechanisms: pre-paying entropic costs through ordering in the ground state and spatial compensation through increases in conformational entropy in distal regions after binding. In examining these mechanisms, it also becomes clear that conformational entropy is highly intertwined with classic definitions of conformational changes. We argue that given the ample evidence of the biological significance of conformational entropy, structurally defining the ensembles encoding conformational entropy will open new paths for control of binding, catalysis, and allostery.

THE PROBLEM OF ENTROPY AND BINDING

During protein folding, proteins go from a disordered linear polymer, which can access almost infinite states[1], to a globular, folded form, where access to most states, especially in backbone atoms, is restricted by steric clashes[2]. Consequently, the protein experiences a substantial reduction in entropy due to a loss of access to a significant number of microstates. The relationship is characterized by the Boltzmann entropy equation ([$S=k_B\ln(W)$]; S=entropy; k_B=Boltzmann's constant; W=number of microstates)[3]. Thus, the transition from an unfolded to folded protein involves a reduction of nearly iso-energetic states, which can also be represented by narrowing the free energy landscape. Even though the free energy landscape narrows upon folding, the basins still have width. The width of the basin defines the accessible states, representing the conformational ensemble of a folded protein (also known as the "native state"). The conformational states accessible through thermal fluctuations play a key role in activity. Not only can they accommodate new functions, but the extent of this native state conformational ensemble also correlates with the protein's evolvability[4]. While different backbone conformations can be populated in the native state of some proteins, side chains populate multiple conformations in all proteins - even in the cores of crystallized proteins[5]. This residual conformational heterogeneity of side chains likely accounts for a significant amount of remaining conformational entropy within the native state[6].

To further illustrate the concept of side chain conformational entropy, consider a single serine side chain in a polypeptide. In the unfolded state, the serine side chain can access all three rotamers[7] equally, and the backbone is only restricted by the neighboring residues on the chain(**Figure 1A**). Upon folding, the serine backbone is restricted by steric clashes to a single conformation[2], with residual harmonic motion about its mean position, incurring a significant entropic penalty. However, the side chain may still have access to multiple conformations (at some minimum occupancy, say ~1%), as this small, slightly polar amino acid can readily interact with different alternative environments. Steric clashes with surrounding amino acids may eliminate access to one rotamer. However, the other two rotamers will still be accessible, leading to a reduction, but not the elimination, of side chain conformational entropy in the folded state (**Figure 1B**).

We can visualize this in a free energy landscape as well. The loss of rotameric states is observed as the loss of that energy basin. The higher energy makes it much less likely for that conformation to be populated. However, the side chain retains conformational entropy by being able to anharmonically move between conformations, represented as the movement over the small barrier separating the energy wells in the free energy landscape. The side chains can also move harmonically about each conformation, represented by the width of each distinct well. The entropy is determined by the relative population of the two conformations and the contributions of the harmonic motion within rotameric conformations.

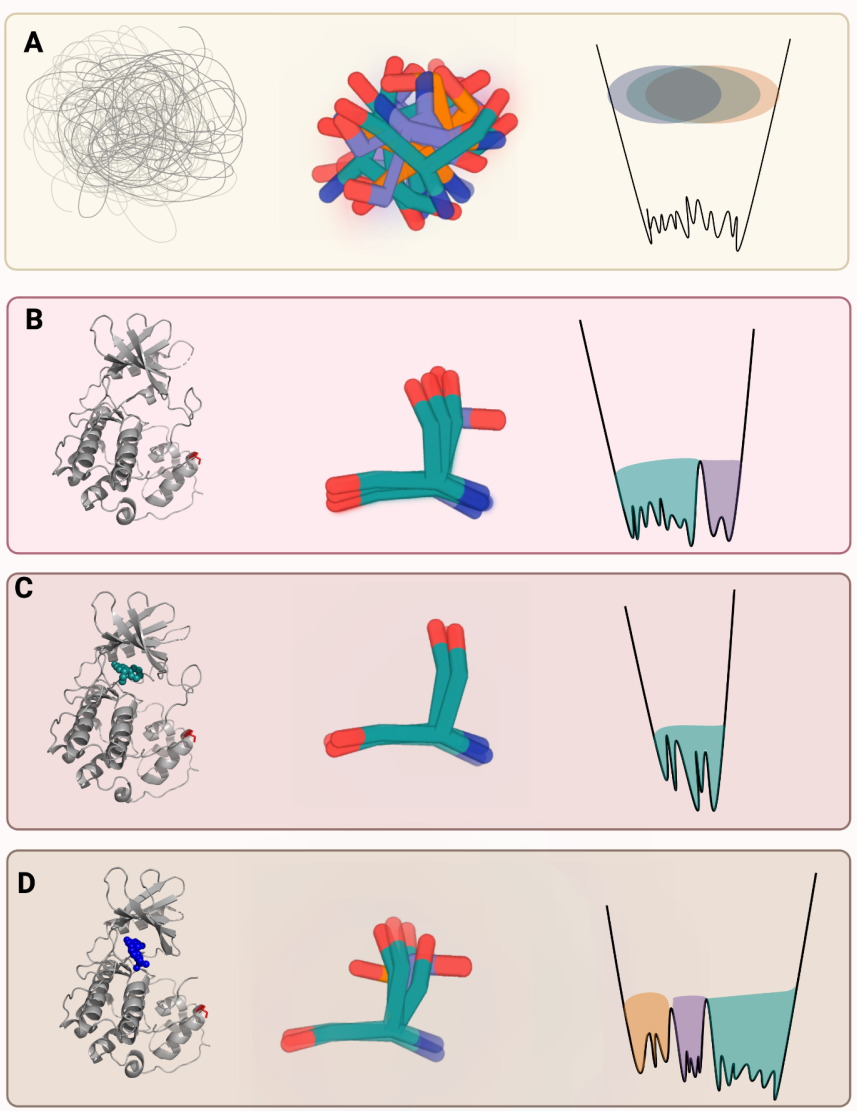

Figure 1. Examples of a estimated conformational entropy for a serine side chain in different contexts. (Left) The structural state of the entire protein with the serine in red. (Middle) The conformation(s) of the serine. (Right) The relative location of each conformation on the energy landscape of the serine side chain.
(A) In an unfolded state, the entire protein and the serine side chain access many different conformations, with each conformation sitting at the top of a wide energy landscape.
(B) Upon folding, the majority of the protein is structured. This results in a large loss of entropy in the serine side chain. The serine can only access two states, separated by an energy barrier. One conformation has significantly more harmonic entropy, as shown by wider energy well (teal), relative to the other conformation (purple). This harmonic motion is modeled by B-factors in X-ray crystallography.
(C) When one ligand binds, the serine loses an additional conformation, reducing its conformational entropy to one state (teal). This is also observed by narrowing the energy landscape well.
(D) However, when a different ligand binds, with different interactions and no direct contact to the serine residue, the serine gains conformational entropy by accessing all three rotamer states, but again with different widths due to differences in harmonic motion between each conformer. This leads to an INCREASE in conformational entropy for this residue upon ligand binding relative to the unbound state (B).

Perturbations to proteins, including ligand binding and mutations, alter the free energy landscape[8]. Returning to the serine example, the ligand may directly interact with the side chain and strongly favor one of the remaining conformations (**Figure 1C**). This type of stabilization is often described as a conformational change, but it is more accurately described as a conformational redistribution. The reduction in conformations from two accessible rotamers to dominantly populating a single rotamer further reduces conformational entropy.

Ligand binding can also, counterintuitively, increase the conformational heterogeneity of protein side chains[9,10]. For example, ligand binding at one site may stabilize alternative conformations of residues surrounding our serine residue. Conformational entropy can increase when this conformational redistribution results in a void that permits populating the third serine rotameric state (**Figure 1D**). Due to the small population difference of this change, the most probable structure may look similar. However, the accumulation of many such population differences can cause a meaningful energetic difference. Such redistributions can increase conformational entropy and provide an allosteric free energy source for protein functions, including binding and catalysis.

This example illustrates the complex interplay between conformational entropy, structure, and function. Structural descriptions of allostery are frequently based on single conformer structures and described in a “domino effect” of conformational changes[11]. However, even in systems dominated by large structural changes, changes in conformational heterogeneity should be considered. Indeed, over 40 years ago, Cooper & Dryden postulated that protein thermal fluctuations, without change in mean atomic positions, were sufficient to provide “dynamic allostery” between sites[12]. The free energy landscape provides a path to reconcile the “domino effect” with theoretical “dynamic allostery” concepts[8]. Our dominant models to explain the energetics of ligand binding and catalysis rely on narrowing the free energy landscape that maximizes complementarity via the precise orientation of receptor/enzyme side chains and backbone atoms. This complementarity incurs an entropic penalty over a less restricted unbound ensemble. However, the entropic penalty that occurs locally for binding site residues may be partially compensated by increased conformational heterogeneity in other parts of the receptor/enzyme, as shown by the toy example in **Figure 1D**. An alternative mechanism would be for the protein to evolve a pre-organized unbound state that more heavily favors the ligand-bound conformation.

Here, we review how conformational entropy modulates the free energy of various macromolecular functions. Proteins have likely evolved to modulate their conformational entropy to reduce the impact of conformational entropy loss during functional events such as binding or catalysis. We believe this occurs through two models, through increases in conformational entropy in spatially distant areas from the functional site or by pre-paying entropic costs through ordering in the ground state. Further, they have evolved to have conformational entropy facilitate larger conformational changes, such as reorganizing a domain for binding or in large macromolecular machines by making transition states more energetically favorable. Therefore,

conformational change and conformational entropy should be through the lens of conformational ensembles and energy landscapes. Emerging examples suggest that the time is right to begin manipulating conformational ensembles prospectively with this view in mind, reaping the benefits of every last Joule of free energy in rational protein and ligand design.

QUANTIFYING CONFORMATIONAL ENTROPY: NMR ORDER PARAMETERS CORRELATE WITH EXPERIMENTAL MEASURES OF BINDING ENTROPY

Nuclear magnetic resonance (NMR) has established the strong connection between protein function and conformational entropy. Through NMR relaxation techniques, we can measure the disorder in specific bond vectors, quanitatively representing this disorder as an order parameter that directly correlates with conformational entropy[13–16]. The resulting order parameters provide a site-specific measurement of the amount of disorder of NH or methyl groups on the picosecond-nanosecond timescale[17,18]. Starting with Calmodulin and extending to other complexes, Wand and colleagues have used methyl side chain order parameters to estimate conformational entropy[14]. Conformational entropy can be approximated from isothermal titration calorimetry after accounting for solvent and rotational-translational degrees of freedom. The remarkable correlation between entropy estimates from methyl order parameters and calorimetry is referred to as the "entropy meter"[19,20].

However, there are many equivalent side chain conformational ensembles that can produce a given order parameter value. For example, increases in disorder can result from jumps between rotamer states or increased harmonic motion within rotameric wells. The single value of the order parameter is, therefore, degenerate with regard to the direction and extent of the different conformations a side chain populates[15]. This degeneracy becomes more complex as order parameters do not detect exchanges occurring on slower timescales. Side chain conformations can be described as a hierarchy of motions, with slowly interconverting macrostates (mostly on the rotamer scale) composed of quickly interconverting microstates (within rotamer wells). Consequently, the relationship between side chain conformational ensembles and changes in conformational entropy is particularly ill-defined, even when considering the relative change in conformational entropy for an individual residue.

It is perhaps unsurprising then that order parameters, or changes in order parameters, are not well correlated with solvent exposure or interatomic interactions[21]. Some of the limitations in these studies are due to the sparsity of the measurements: only very few proteins have had methyl order parameters measured, and only side chain methyls are measured! More coherent trends may emerge if disorder across all residue types can be measured. Nonetheless, the lack of connection between this quantitative measurement and any conformational ensemble model significantly diminishes our understanding of the mechanisms by which conformational entropy modulates free energy.

Ideally, molecular dynamics (MD) simulations would allow us to observe the fluctuations probed

by NMR directly[22]. Indeed, Sharp and Wand suggest that even short and unconverged simulations can retain a reasonable correlation to observed order parameters and conclude that methyls may be sufficiently well dispersed to be a proxy for total side chain entropy[15]. However, it is clear that significant protein dynamics, with ramifications for conformational entropy, occur on the ms-s timescale[23,24]. Most specialized hardware and acceleration schemes are only beginning to reach those timescales. Thus, while MD can provide insight into the interplay of individual atomic motion over time, our ability to use it to derive insights about conformational entropy is still limited by both the accuracy of force fields and computational costs for accessing long time scales[25].

ENSEMBLE MODELING OF X-RAY AND CRYOEM CAN QUANTIFY CONFORMATIONAL ENTROPY

Alternatively, a readily available source of insights on conformational entropy emerges from ensemble modeling of X-ray crystallography and CryoEM data. These experiments collect ensemble-averaged data from tens of thousands to billions of individual molecules, capturing all conformations regardless of the timescale of exchange. To model the underlying conformational ensemble, careful attention must be paid to the noise arising from radiation damage and poor alignment, which manifests in loss of diffraction in crystallography[26] or inability to generate a high-resolution reconstruction in single particle EM[27]. Of course, genuine conformational heterogeneity also leads to a loss of alignment. However, excitingly, the heterogeneity of protein side chains and loops is often not disruptive to diffraction or alignment. For high (better than 2.0Å) resolution datasets, statistically significant signals in the density map remain unaccounted for by a single PDB coordinate set[28]. Side chain and backbone atoms can be modeled into discrete alternative conformers, leaving B-factors to approximate the harmonic motion around each conformer[29]. These ensemble or multiconformer models better represent the underlying conformational landscape.

Such high-resolution density maps can therefore provide similar information to NMR order parameters because they reveal rotameric jumps as alternative conformers and harmonic motions as B-factors. By determining the distance between each alternative conformation and the weighted B-factor of each conformer, a "crystallographic order parameter" can be obtained that agrees reasonably with NMR relaxation experiments[30]. Importantly, due to the time-averaged nature of diffraction or single particle imaging, this metric goes beyond the pico-nanosecond timescale of NMR order parameters nor is it limited to methyl-containing side chains or torsion position.

By building multiconformer models, which provide changes in conformational distributions, and calculating crystallographic order parameters, which provide changes in the magnitude of conformational heterogeneity, we have the potential to structurally rationalize changes in conformational entropy to changes in protein function. Recently, we built multiconformer models for ~1200 cryogenic X-ray crystallography datasets to interrogate the role of conformational

entropy in a widespread and fundamental protein function: small molecule ligand binding[9]. A simple model for the relationship between ligand binding and conformational entropy is that proteins pay an entropic penalty by ordering the binding site to facilitate stable (enthalpic) interactions with the ligand. In addition to the free energy gained by these new enthalpic interactions, the displacement of water molecules from the binding site provides an additional (entropic) free energy source. However, it was unclear how the rest of the protein changed in response to the ligand binding. It's conceivable that the entire protein could become more rigid, leading to a further reduction in entropy. Alternatively, the remaining parts of the protein might maintain their existing flexibility, or even experience an increase in flexibility, increasing entropy.

After building multiconformer models, we compared the crystallographic order parameters between ~600 paired ligand-bound and unbound structures. In line with expectations, side chains in binding sites tended to become more ordered when bound to a ligand. While this conforms to our intuition about binding sites, there are a range of behaviors observed in binding sites, including examples where the binding site becomes more flexible. We also observed that ligands with fewer hydrogen bonds or higher logP, indicative of more greasy or hydrophobic ligands, tended to have a reduced entropic penalties upon binding. This result aligns with the idea that interactions with a more specific distance and orientation dependence, like hydrogen bonds, can lead to a greater loss of conformational entropy.

While our results indicated that ligand binding incurs a conformational entropy penalty locally at the binding site, we also discovered a potential entropic mechanism that favors binding: residues that are distant from the binding site tend to become more flexible upon ligand binding. Moreover, the magnitude of this increase inversely correlated with how much flexibility was lost around the binding site. This inverse correlation suggested that in addition to new interactions and the release of water molecules, proteins often compensate for the entropic penalty at the binding site by increases in conformational entropy elsewhere in the protein. Considering the molecular origin of these changes returns us to the example of the Serine side chain above. Local ordering at the binding site tends to reduce conformational entropy (compare **Figure 1B** and **C**), whereas many non-contacting residues can increase conformational entropy (compare **Fig 1B** and **D**).

We now think of this trend as entropic “spatial compensation” - a penalty locally for ordering is compensated by increased disorder at distant sites. Correlation is, obviously, not causation - and there are many outlier cases to the overall trend. However, the spread of changes at distant sites suggests that conformational entropy is likely modulated by evolution and (perhaps unwittingly) medicinal chemistry. While our study represented the largest exploration of how conformational entropy impacts ligand binding, the major findings agree with the themes emerging from NMR studies, which have also suggested cases of entropic “spatial compensation” upon ligand binding. Further, entropic spatial compensation can extend beyond ligand binding to many other macromolecular functions. We argue that an atomistic explanation of the mechanisms of conformational entropy, which could be obtained from X-ray

crystallography, CryoEM, or long-timescale simulations, will accelerate our ability to rationally control this aspect of the free energy driving biological function.

CONFORMATIONAL ENTROPY GETS REDISTRIBUTED DURING BINDING, MODULATING AFFINITY

The correlation between binding entropy estimated by NMR and measured by calorimetry was first observed in calmodulin, which binds a wide variety of peptides and is allosterically regulated by calcium binding[19,31]. Upon binding a peptide derived from smooth muscle myosin light chain kinase peptide (smMLCKp), NMR methyl order parameters across calmodulin increased (indicating a loss of conformational entropy). This trend was especially prominent around the smMLCKp binding site. However, the expansion of these studies to six peptides with roughly the same affinity to calmodulin, revealed various changes in entropy with no clear discernable spatial pattern. Due to the similar size and composition of the peptides, the entropic changes from water displacement and the ligand degrees of freedom are nearly identical, leading to the idea that the changes resulted from protein conformational entropy. After accounting for the solvent and ligand effects, the residual conformational entropy measurements from calorimetry could be favorable, disfavorable, or neutral contributions to the overall binding free energy. Remarkably, the NMR order parameter change upon peptide binding was correlated with direction and magnitude of this calorimetric residual conformational entropy, allowing the calibration of the "entropy meter"[19,31].

Kalodimos and colleagues extended the entropy meter concept with a series of studies on the catabolite activator protein (CAP), a dimeric protein that binds DNA in a cAMP-dependent manner[32,33]. First, they measured NMR order parameters showing that cAMP binding causes a local rigidification for the first binding event and a much larger and more global rigidification when the second cAMP binds. This pattern is paralleled in the ITC measurements which show that entropy disfavors the first binding event and that the entropic penalty is even stronger for the second cAMP binding event. In subsequent papers, they used a series of mutants that are distant from the DNA-binding interface that have markedly different affinities for DNA[34] driven by difference in entropy spanning a range of ~+/-20kcals. The calorimetry entropy measurements were highly correlated with the NMR entropy measurement. Because the DNA ligand is constant through the experiments, we can be even more confident that solvent and ligand contributions to entropy are minimally perturbed than when considering the related peptides in the Calmodulin analyses.

Building on these results, a later analysis by Wand and colleagues examined the relationship between protein binding entropy and conformational entropy in 28 protein-ligand complexes involving eight different proteins (including Calmodulin and CAP, and also galectin, lysozyme, PDZ, DHFR, SAP, HBP(D24R))[20]. Even more remarkably, the correlation of the "entropy meter" and the residual entropy from ITC holds across this diverse group of proteins. However, how

conformational entropy varies across these complexes has resisted a causal mechanism[35,36].

An instructive example of how conformational entropy impacts binding affinity comes from Lee and colleagues in their study of the third PDZ domain from PSD-95/SAP90 (PDZ3). The truncation of the alpha helix 3 (α3) of PDZ3 reduces peptide binding affinity by 21-fold, even though it is over 6 Å away from the binding site and makes no direct contacts with the ligand[37]. Moreover, the lack of chemical shift perturbations in the rest of the protein upon α3 truncation indicates that it does not greatly change the average structure at the binding site. Rather than a detectable change of structure, they observed differences in side chain flexibility throughout the protein measured by NMR order parameters. While the dynamics are nearly identical in the bound state, the unbound states differ drastically. In the unbound state, the side chains are much more rigid in the WT PDZ3 than in the truncated version. This results in a greater difference in side chain flexibility in the truncated variant, compared to WT. Therefore, the truncated variant pays a much larger entropic penalty upon binding peptide because its apo state has more entropy. This trend is also observed in the ITC data (**Figure 2**).

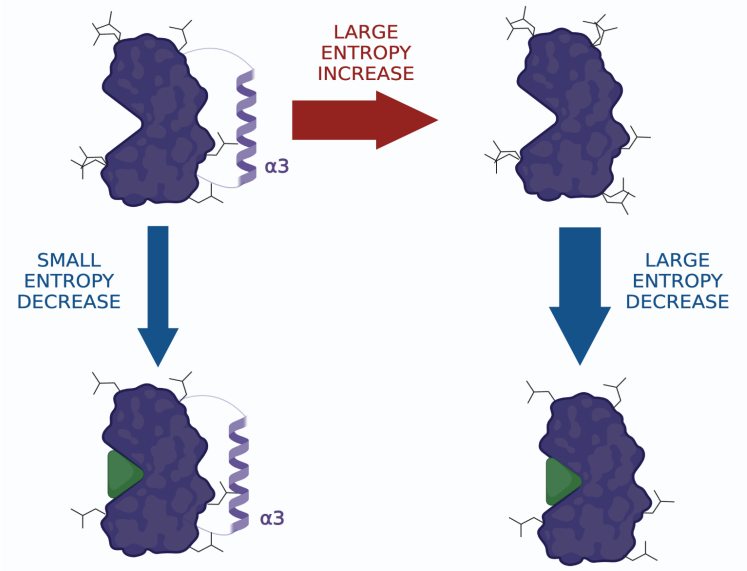

Figure 2. When wildtype PDZ3 (top left) has the α3 helix truncated (top right), there is a large increase in conformational entropy (represented by the number of leucine side chains conformations). Both the wildtype and α3 truncated PDZ3 muxst pay an entropic penalty in order to bind it's ligand (green triangle). Because of their differences in entropic starting points (α3 truncated having higher entropy in the unliganded state), the α3 truncated pays a much larger penalty, helping to drive the 21-fold lower binding affinity in the α3 truncated compared to wildtype.

This example illustrates the significance of conformational entropy in modulating ligand binding affinity. However, the mechanistic explanation for changes in "side chain flexibility" remain unsatisfying. Clearly, chemical shift perturbations and analyses of static crystal structure are not sufficient to derive insights into the types of packing changes that change order parameters. We infer that local packing changes upon deletion of α3 lead to distant changes in packing; however, it is unclear to what extent those changes increase harmonic motion of residues in

their current rotamers or enable access to new conformations. Without satisfying answers to those questions from multiconformer X-ray data, distinct NMR experiments, or long timescale MD, the pattern of disrupted hydrogen bonds, salt bridges, and hydrophobic contacts cannot be discerned or manipulated. It also reveals the false dichotomy between “structure” and “dynamics”. Even subtle changes in structure, below the limit of detection from NMR chemical shift perturbations or comparison of moderate resolution crystal structure, can alter side chain dynamics of neighboring residues. These alterations can have highly non-linear patterns as the ensemble is redistributed away from the perturbation site. The result is a significant thermodynamic effect driven by entropy. Although the α3 truncation is an artificial perturbation, PDZ domains are normally components of multidomain proteins that can be regulated by phosphorylation regions that can reversibly bind to the PDZ domain like α3, suggesting that the mechanism described above is more than a thermodynamic curiosity[38].

CONFORMATIONAL ENTROPY IS INTERTWINED WITH ALLOSTERY TO DRIVE BINDING COOPERTIVITY

Ligand binding cooperativity is at the heart of allostery. The simplest cases are those of identical binding sites in multimeric proteins. Consider a dimeric enzyme. Positive cooperativity occurs when the affinity at a second site is enhanced when the first site is occupied. Entropy can play a key role in modulating the differences in the apparent affinities between the two sites. For example, binding at the first site can order the side chains locally at the binding site, incurring an entropic penalty. Any change in the structure or dynamics of the second site promoted by the first binding event provides a mechanism to alter the apparent affinity of the second binding event. A simple case of positive cooperativity can be envisioned where the ordering at the first site also stabilizes the ordered (binding competent) conformations at the second site. Therefore, the ordering at the second site is considered part of the entropic penalty of the first binding event, reducing the entropic penalty of binding at the second (otherwise equivalent) site. This leads to an INCREASE in affinity as the entropy penalty is lower, but the enthalpic contribution are the same as in the first site. Negative cooperativity can occur through the opposite mechanism: ordering at the first site leads to destabilization of conformations consistent with binding at the second site, increasing the entropic penalty of binding. These models extend the “reorganization energy” in classic models of allostery, which invoke static conformational changes, into the realm of conformational ensembles[39].

In addition to the CAP outlined above, entropically-driven cooperativity has also been implicated in CzrA (chromosomal zinc-regulated repressor), a DNA-binding protein. CzrA demonstrates an altered DNA binding affinity due to the presence of a zinc ion, which is located more than 15 angstroms away from the DNA binding site. Without Zn binding, CzrA increases side chain flexibility upon binding DNA, specifically in areas distant from DNA. However, upon Zn binding, side chain flexibility decreases throughout the protein, resulting in a lower DNA binding affinity[40]. This example demonstrates how multiple states that are compatible with binding can

entropically influence the free energy of binding.

A beautiful example of small molecule example of positive cooperativity through changes in side chain dynamics has been uncovered in human thymidylate synthase (hTS) by the Lee group[41,42]. This dimeric enzyme catalyzes the conversion of deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP). Careful calorimetric measurements at multiple protein concentrations demonstrated that dUMP binding has positive cooperativity between sites. Interestingly, the entropic component reduces affinity by ~2.5 kcal/mol for the first binding event. The free energy of binding is favorable (~-6kcal/mol) because of a strong enthalpic component (~-8.5 kcal/mol). In contrast, entropy favors the second binding event by -2.5 kcal/mol, and there is a relatively weaker enthalpic component (~-5 kcal/mol). The cooperativity results from a change in free energy (~-1.5kcal/mol) across the two binding sites. What is the origin of this ~5kcal/mol swing in entropy? Their initial investigation suggested side chain conformational entropy, but they were not able to rule out additional contributions to changes to ordered solvent or backbone flexibility[42].

Subsequent NMR studies revealed a panoply of dynamic changes across timescales. Relaxation dispersion and Chemical Exchange Saturation Transfer (CEST) experiments revealed multiple states with small populations (~1%) that exchange on slower time scales, which are suppressed by ligand binding. Residual Dipolar Coupling (RDC) data confirmed that the dominant state populated in the apo form resembles the active, ligand-bound conformation, not a previously crystallized inactive conformation. Therefore, at least at the level of precision of RDCs, gross structural changes cannot explain the huge entropy swing. Side chain dynamics experiments interpreted through the “entropy meter” revealed a ~10kcal entropic penalty as the enzyme moves from both sites apo to both sites dUMP-bound. The loss of slow dynamics and the dramatic change in side chain dynamics can be reconciled through the view of population shuffling[42,43], which explains how such changes can be linked across timescales. Interestingly, much of this rigidification occurs around the binding site of the cosubstrate mTHF, presumably preparing the cost of the co-substrate binding.

The emerging model is that the first binding event significantly reduces side chain conformational heterogeneity across the entire protein, thus pre-paying the entropic penalty of the second binding event (**Figure 3A**). This model suggests that almost all protein conformational entropy is reduced during the first binding event, leading to virtually no change in protein conformational entropy in the second. While the model does differ from the calorimetry results (calorimetry: ~5kcal/mol, entropy meter: ~10kcal/mol), it's still a viable hypothesis. This is because NMR measurements can't be conducted with a single dUMP bound to dimeric hTS, making it impossible to ascertain the intermediate entropy meter reading. Further, solvent entropy may play a large role here, as observed in the homodimeric enzyme fluoroacetate dehalogenase during catalysis[44].

Remarkably, the role of entropy in driving positive cooperativity is highly specific to hTS bound

to the substrate dUMP. hTS's substrate (dUMP) and product (dTMP) differ by only a single methyl group, provided by the co-substrate, N5, N10-methylene tetrahydrofolate. However, no positive cooperativity is observed by calorimetry for dTMP binding. Chemical shift perturbations between hTS bound to dUMP and dTMP indicate similar patterns, suggesting no substantial structural changes. However, in agreement with the calorimetry data, the changes in side chain order parameters between hTS-TMP and apo hTS are negligible. These results point to the limitations of static structures and chemical shift changes to infer changes in the conformational ensemble. The large difference in conformational heterogeneity between dUMP and dTMP, with the corresponding large energetic effect, is difficult to reconcile with a chemical change as small as a methyl group.

A key to understanding how this "magic methyl" could have such a large effect may result from the difference in how the N-terminal disordered segment of hTS interacts with the globular part of the enzyme. Bonin et al hypothesized that transient contacts made by the disordered segment are sensitive to the chemical identity of the ligand and the subtle conformational changes each stabilizes. Rather than suggesting a disordered-ordered transition for this segment (and the entropic costs associated with such a transition), they suggest that the pattern of transient contacts between the disordered segment and the globular region can alter side chain dynamics. Side chain dynamics are quenched, due to contacts with the disordered segment, incurring an entropic penalty for dUMP binding, but are relatively unchanged for dTMP binding. Intriguingly, this segment is not present in the *E coli* homolog of TS, which is also dimeric but shows no cooperative behavior[45], suggesting the modulation of thermodynamics by side chain entropy can be a powerful mechanism exploited through evolution.

Whether it is through interaction with the N-terminus or other residues in the binding site, the addition of the methyl group restricts the accessibility of surrounding residues to additional conformational states, leading to a decrease in conformational entropy. This problem is general: similar ligands generally bind with similar binding poses and make similar interactions in the binding site. The similarity in these factors suggest that specificity is unlikely to be solely dictated by enthalpic interactions. Alterations in conformational entropy, created by generating or eliminating minor voids around the binding pocket, can potentially have profound energetic consequences for binding specificity. Examining the impact of conformational entropy in a ligand series, might identify chemical attributes that impact protein conformational entropy. Revealing the diverse ways in which various ligands affect conformational entropy carries profound implications, potentially paving the way for a novel approach to rational drug design. While these examples point out how pre-paying entropic costs is a route to positive cooperativity, it is also possible that conformational entropy could increase upon all binding events. In that scenario, conformational entropy would more directly drive positive cooperativity.

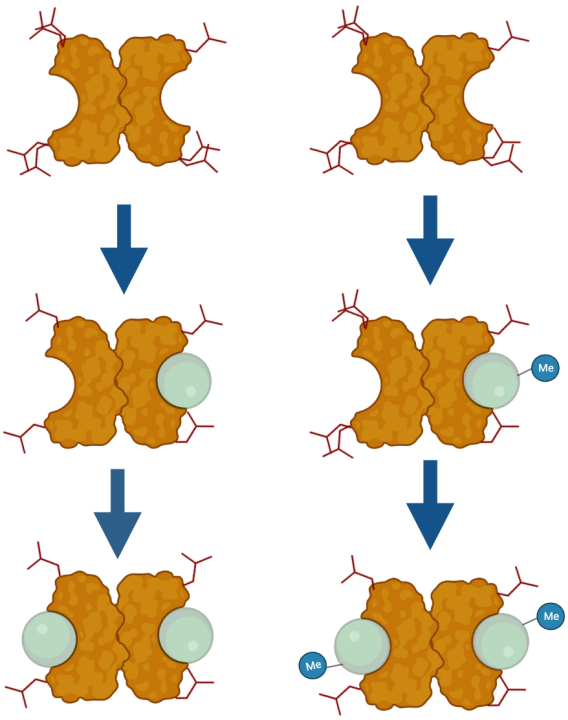

Figure 3. (**A**) Human thymidylate synthase (hTS) displays positive cooperativity when binding its substrate dUMP (yellow). This is driven by entropy, with the binding of the entropic contribution of the first dUMP hurting the binding affinity (+2.5 kcal/mol), and the entropic contribution of the second dUMP helping the binding affinity (-2.5kcal/mol). Using NMR, we know that binding both dUMP molecules leads to a widespread reduction in conformational entropy. Based on the calorimetry measurements and NMR data, this leaves us with the hypothesis that the majority of the reduction of conformational entropy is upon the first binding event.
(**B**) When hTS binds to its product dTMP (burnt orange), there is no positive cooperativity observed (both binding events have similar binding affinities). Additionally, the entropy in both binding events hurts the binding affinity.

CONFORMATIONAL ENTROPY DRIVES ALLOSTERY AND COOPERTIVITY IN MONOMERIC PROTEINS

While the CAP, CzrA, and hTS examples detail entropic contributions to allostery in multimeric proteins, entropy-driven binding cooperativity has also been observed in monomeric proteins, particularly in protein kinases [46–48]. For example, in cyclic adenosine 3′,5′-monophosphate (cAMP)–dependent protein kinase A (PKA-C), entropically driven cooperativity is observed between the nucleotide and substrate binding sites[46]. Inhibitors that more resemble the natural substrate, ATP, lead to ordering at the substrate binding site, highlighting how evolution (and medicinal chemistry efforts, perhaps unwittingly) have manipulated conformational entropy. A greater structural understanding of these events may have significant relevance to double-drugging strategies[49,50].

Entropy can also play a role in allostery of membrane proteins. Such allostery can control, as observed in GPCRs, downstream signaling consequences [51]. While GPCR extracellular

orthosteric site binding agonists and antagonists can exhibit similar binding affinities[52], antagonists stabilize an "inactive" state and agonists favor a broader "active" ensemble[53]. This active ensemble facilitates intracellular interactions with different G proteins or arrestins. Interestingly, different agonists can lead to subtle differences in their “active” ensemble, resulting in preferential interactions with different partners and activation of distinct signaling pathways. Multiple studies have identified the importance of side chain conformational dynamics in connecting the logic of which orthosteric binders bias binding at the allosteric site[54,55], suggesting a role for conformational entropy in allostery.

In addition, the theme of “spatial compensation” may be operative in GPCRs as well: subtle changes in conformational ensembles of side chains outside of the allosteric networks linking the orthosteric and allosteric sites may also influence the relative stability of specific conformations compatible with distinct binding partners. Leveraging conformational entropy to alter protein-protein interaction preferences may be a common and particularly important phenomenon in GPCRs because their apo states have high conformational entropy, to offset the entropic penalty of sitting in the membrane[56].

COVALENT MODIFICATIONS CAN SHIFT THE CONFORMATIONAL ENSEMBLE VIA CONFORMATIONAL ENTROPY

Thus far, we have outlined how conformational entropy controls reversible ligand binding. Covalent changes, such as protein post-translational modifications (PTMs) and mutations can also change the conformational ensemble. While the dominant mode of thinking about the effect of PTMs is through stabilizing large conformational changes, changes in conformational entropy can significantly modulate the energetics of the modified ensemble. Two relevant examples from bacterial response regulators show how phosphorylation can bias flexibility and downstream signaling. In CheY, autophosphorylation results in stabilization of conformations that are compatible with motor proteins controlling chemotaxis. While a decrease in flexibility is observed by NMR around the motor protein binding site, this entropic cost is partially recovered by increased side chain flexibility, particularly along the pathway between the autophosphorylation and the motor protein binding site[23].

Similar phenomena have also been observed in a related bacterial response regulator, Nitrogen Regulatory Protein C (NtrC), with particularly interesting connections between increases in side chain flexibility and an entropically lowered barrier of the transition state between conformations[57]. These results further support a connection between changes in fast dynamics lending a hand for motions that occur on a slower timescale. This drives the connection between conformational entropy and traditional conformational changes. Conformational entropy can help decrease the energetic differences or barriers for motions that occur on a slower timescale[58]. Other PTMs likely modify conformational ensembles in similar ways through the covalent connection to the protein.

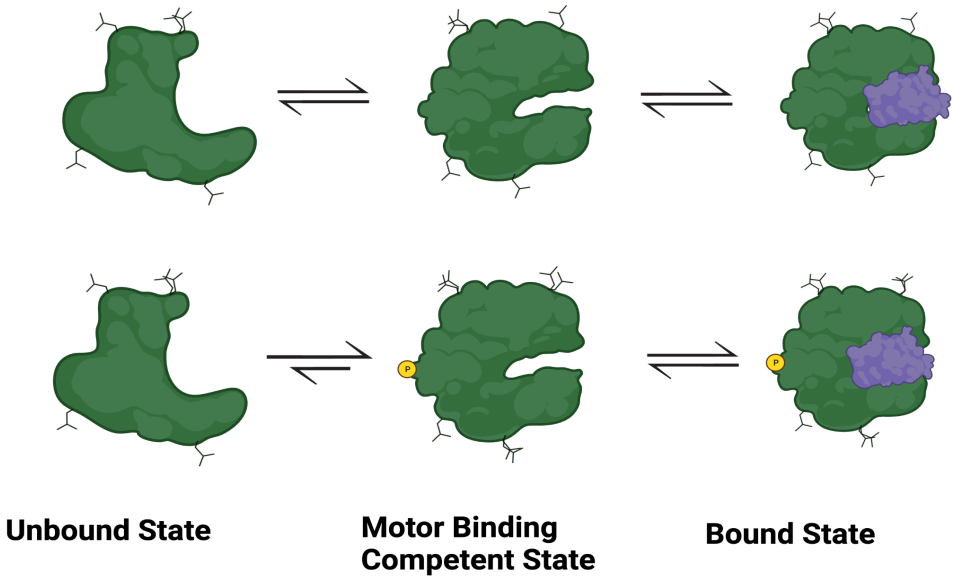

Figure 4. CheY must undergo a large conformational change from it's apo state to it motor protein bound state. Autophosphorylation increases the conformational entropy, increasing the probability of the motor binding protein state.

Mutations can be thought of similarly to PTMs in that they alter the covalent composition of a protein. It is now commonplace to rationalize the effects of active site mutations in the context of disease or engineering based on structural models. While the list of disease-causing mutations is consistently expanding, the functional implications of most mutations, especially those distal from functional sites, remain poorly understood[59,60]. Similar challenges emerge when analyzing mutations emerging from directed evolution campaigns. Broadly, most mutations away from functional sites are interpreted with regard to folding stability or at best a three state model (folded-functional, folded-non-functional, unfolded). As outlined in the beginning of this review, folding is associated with a large entropic penalty, but mutations that increase side chain entropy can help repay some of this debt[61].

An example of how side chain entropy can influence folding stability emerged from an experimental selection scheme for stabilized variants of Chymotrypsin Inhibitor 2 (CI2). This work identified a set of substitutions with interesting underlying differences in stability. The introduction of a single hydrophobic core mutation (L49I) results in a less stable folded state, whereas an alternate hydrophobic core mutation (I57V) enhances folded state stability. The combination (L49I/I57V) showed significant epistasis and was more stabilizing than I57V alone[62]. There are a few potential models for reconciling these results. The simplest model is that the stabilizing mutant has more perfect hydrophobic packing, due to more ordered side chains. Conversely the destabilizing mutant would create packing defects associated with side chain disorder. It would follow that the double mutant is more stable because of more perfect

complementarity packing in the hydrophobic core and thus further reduced disorder. However, the results from NMR order parameter measurements are the exact opposite of this model! The destabilizing L49I variant correlated with more ordered side chains and the stabilized mutants had increased dynamics[63] (**Figure 5**). Again, the entropy meter concept provided a rough quantification that correlated, but overestimated, the degree of stabilization. The overestimate may reflect some entropic costs that are not accounted for, such as internal void volumes.

The authors hypotheszied that the inconsistency between the entropy meter and degree of stablization is due to the NMR focus on methyl side chains. To investigate this, the authors employed MD simulations to complement NMR and crystallography[63]. First, they ran multiple independent simulations seeded from alternative conformations of the side chains resolved in the crystal data. These alternative conformations are likely separated by a significant barrier and therefore this helps sample conformations representative of longer timescales. Second, they reweighted the MD data (with rigorous overfitting checks) to provide better overall agreement with the NMR observables. While a complete explanation for the inconsistency between the stability measurement and entropy meter did not emerge, the two MD methodological improvements point the way forward to reconcile NMR with a structural rationale. It also emphasizes the need for more structural rationale to identify the underlying cause of the overestimate of stability.

Additional examples of mutations that increase flexibility beyond side chain dynamics being correlated with the relative stability of different states are emerging. In Adenylate kinase, mutations that affect local unfolding, as measured by differential scanning calorimetry and NMR, lead to favorable catalytic properties through entropic stabilization[64]. Neutron scattering, which reveals average changes in flexibility, suggested that an engineered creatinase is stablized, relative to wildtype, by increased entropy in the folded state[65].

Mutations involving the SARS-CoV-2 spike protein have taken on a prominent place in the public consciousness and illustrate how entropy can be tuned in either direction to optimize free energy. The alpha variant mutation, N501Y, which leads to enhanced infection and transmission[66] through increasing the affinity of the RBD to angiotensin-converting enzyme 2 (ACE2). This mutant does not optimze packing at the interface, but increases flexibility of the complex in a way that favors binding entropically[67,68]. In agreement with more traditional concepts of affinity maturation, nanobodies selected to neutralize spike exhibit decreases in flexibility and more optimal interfaces as affinity increases[69]. These examples illustrate the wide ranging effects mutations can have on the relative stability of folded, binding competent, and unfolded states and highlight the need to examine the ensemble effects of mutations. By examining the relationship between mutations and conformational entropy, we can shed light on the implications of disease mutations, evolution's modulation of function, and improve the design of functional proteins.

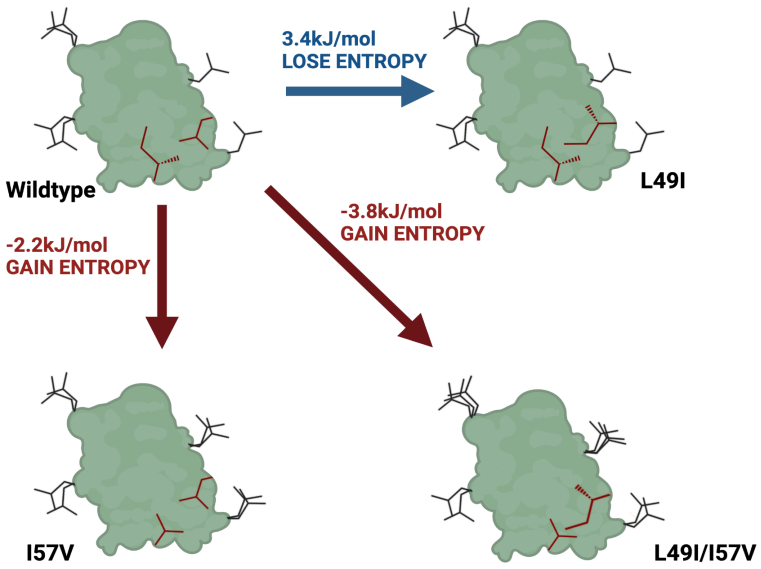

Figure 5. Upon L49I mutation, CI2 becomes less stable (losing 3.4kJ/mol), by losing conformational entropy (represented by the number of 'side chain' conformations in each state) compared to wildtype. However, upon I57V mutation, CI2 becomes more stable due to gaining conformational entropy compared to wildtype. With both mutations, there is even larger gain in conformational entropy compared to wildtype, leading to the highest stability versions between the mutants and wildtype.

DEGENERATE SCAFFOLDS COUNTERBALANCE THE ENTROPIC COSTS OF AN ORGANIZED ACTIVE SITE

Enzymes require the precise arrangement of active site residues for efficient catalysis of reactions, leading to an extreme reduction of conformational entropy of these residues. Despite this, the enzyme can access numerous conformational states, provided these states are compatible with the position and subtle flexibility required by catalytic residues throughout the catalytic cycle[70,71].

The ability for the non-catalytic residues to fluctuate between many states could mitigate some of the entropic cost associated with the exact positioning of the catalytic residues. The concept of pre-paying entropic costs is not new to the enzymatic field. The "Circe effect" hypothesis postulates that the entropic penalty associated with a transition state having little entropy is 'pre-paid' upon substrate binding. This effect allows the enzymatic reaction to pay a much lower entropic penalty, similar to the binding cooperativity discussed above. However, a growing understanding intimates that the enzyme's intrinsic capacity to traverse various states during catalysis also significantly alleviates the entropic component of the activation barrier (**Figure 6**). Enzymes can compensate for the extreme loss of conformational entropy in catalytic residues by increasing it elsewhere in the other protomer. One can imagine that this phenomenon could also occur in a monomeric enzyme by residues distal to the catalytic site increasing their dynamics, as well as a release of water molecules. Understanding how the rest of the protein's conformational entropy supports the entropic penalty of binding site residues is crucial for

comprehending enzyme function.

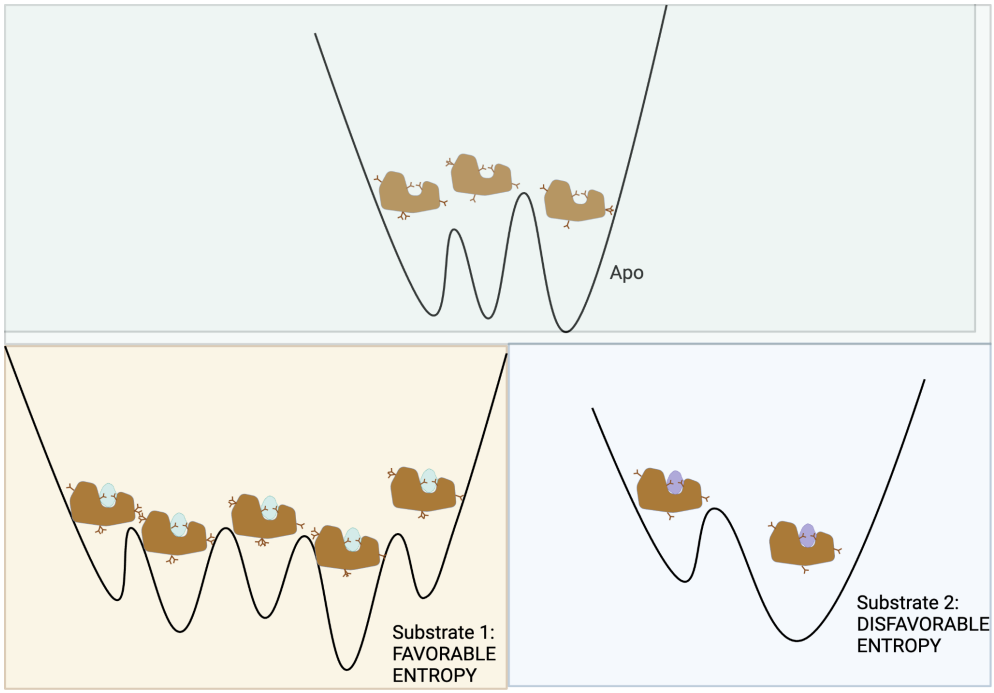

Figure 6. Enzymes can have multiple states compatible with the required position of catalytic residues, increasing the entropy of catalysis, thus lowering the energetics. In the apo form the enzyme has 3 accessible states. When substrate 1 binds (left), there is an increase in states, all while keeping hte catalytic residues positioned correctly, INCREASING the entropic contribution. When substrate 2 binds (right), there is a decrease in states, decreasing the entropic contribution.

A protein's intrinsic capacity to access many states during a transition proces may also be expanded for understanding the energetics of macromolecular machines. These complexes must harness chemical energy of nucleotide hydrolysis to drive DNA replication, RNA transcription, and protein synthesis. A beautiful example of macromolecular machine harnessing conformational entropy comes from the ribosome. The ribosome undergoes many large, enthalpically expensive structural rearrangements transitioning from initiation to elongation. But many components of the ribosomal machinery, including tRNAs, induce thermal fluctuations in the rest of the ribosome, suggesting a role for entropy in understanding the relevant energetics[72]. Using multiple temperature single-molecule fluorescence resonance energy transfer (smFRET), entropy was shown to compensate, almost exactly, for the enthalpic transition penalties between functional states of the ribosome[24]. Even more amazingly, $tRNA^{Met}$ specific entropic changes favored transition towards the initiation state, whereas the $tRNA^{Phe}$ specific entropic changes favored the elongation state. This indicates that the ribosome has evolved to manipulate entropy to favor the state compatible with the tRNA (**Figure 7**). This, and other studies demonstrating how the ribosome harnesses conformational entropy for peptide bond formation[73], draws a picture where the ribosome has evolved to rely on conformational entropy to absolve it of necessary energetic penalties. As we gain more clarity into the catalytic cycle that the ribosome undergoes during translation[74], it is important to consider how the

fluctuations in protein, RNA, water, and metal ions, can assist in the transition between conformations observable by structural biology methods.

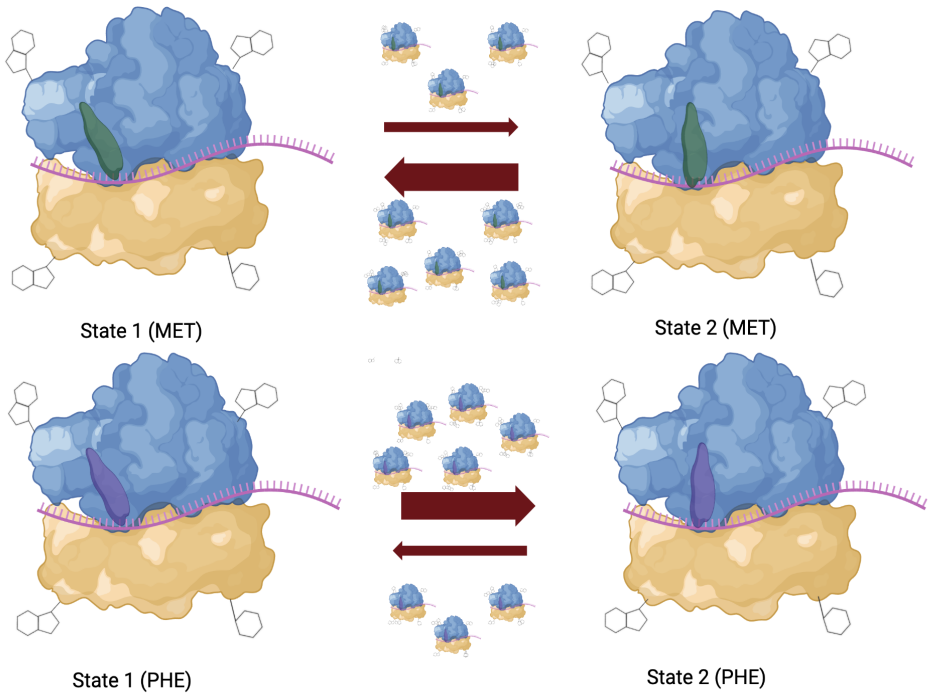

Figure 7. The ribosome has evolved to take advantage of entropy. The ribosome (large subunit-blue; small subunit-yellow, mRNA-red) goes through large conformation changes between the initiation state (state 1) and elongation state (state 2). The enthalpic penalty paid by each state is balanced out by the amount of conformational entropy, represented by the size of the blue arrows the ribosome can access during the transition states. The more conformations the ribosome can access (represented by the number of ribosome states), the higher the conformational entropy. Entropy also helps tipping the balance in tRNAmet (purple) bound ribosome towards the initiation phase and tRNAphe (green) towards the elongation phase.

DISCUSSION

Here, we have reviewed evidence for the impact of conformational entropy on protein function. We identify two primary mechanisms—entropic pre-payment and spatial compensation—as entropic modulators of the change in free energies associated with function. It is clear that an expanded view of the energetic consequences of conformational entropy can reconcile traditional models of allostery with those that invoke “dynamics without conformational change”. Most of this review is focused on the conformational entropy contained within ordered regions, and particularly side chains, of proteins. However, additional important entropic contributions come from intrinsically disordered regions (IDRs), ligands, and solvent[75,76,77].

IDRs can serve as an important way to modulate free energy, often acting as an ‘entropic reservoir’ by increasing the relative disorder upon a binding event as uncovered for human UDP-α-D-glucose-6-dehydrogenase (UGDH)[75]. The equilibrium of the globular domain of UGDH can be shifted to a higher affinity substrate state, by the presence and length of its disordered C-

terminal tail. The length dependence of this effect suggests that the physical origin of this shift lies in the conformational sampling restrictions between the globular domain and the IDR segment. Recent simulation work has articulated the principles of how such forces can be generated by the geometric restrictions globular domains enforce on their attached intrinsically disordered tails[78]. In addition, IDRs can also interact with each other in a manner where both partners retain high disorder and high affinity, as observed between histone H1 and its nuclear chaperone prothymosin-α [79]. This complex has a "fuzzy" character, with many energetically equivalent binding poses increasing the multiplicity of the assembled state. Similar complexes are likely quite prevalent in eukaryotic cells, where a large portion of the proteome contains disordered segments[80,81]. This aspect of entropic free energy may therefore have profound consequences for decoding cell signaling.

The idea of "fuzzy" interactions can be expanded to protein-small molecule interactions. While small molecules lose a significant amount of entropy upon binding, it is likely that many retain a small amount of torsional flexibility, increasing the multiplicity of the complex and thus the entropy.
Modeling multiple ligand conformations frequently alleviates strain and increases agreement with underlying density data, suggesting that many protein-ligand complexes have multiple binding states[82,83,84] and that these likely entropically contribute to binding affinity. Recent examples have demonstrated that optimizing molecules based on their multiple binding states can significantly enhance binding affinity[85,86], highlighting the potential to apply this approach for more targets.

Water can be thought of as a very high concentration and simple small molecule ligand. There is a long literature of how 'release' of a water molecule into bulk solvent increases entropy. To better resolve the mechanisms of solvent entropy, we must improve our treatment of solvent molecules, especially those that are exchanging with solvent and therefore are 'partially occupied', in X-ray and EM model refinement. Purely computational approaches like Watermap and other molecular dynamics (MD) water mapping techniques have shown that many water molecules around the protein and ligand are often partially occupied[87], suggesting that they can have a significant impact on the system's entropy. Currently, the entropy meter evaluates solvent entropy by measuring changes in heat capacity contributions within both non-polar and polar regions of the solvent-accessible surface area (SASA)[88]. While the SASA method works well, it provides little atomic level understanding of the origins of the energies of water molecules. Creating a more atomistic model of the entropic contributions of water molecules moving into solvent may reconcile differences observed with the entropy meter and move us closer to predicting the entropic impact of a perturbation.

The convergence of diverse entropic elements complicates the task of isolating their individual effects[21,89]. It is crucial to develop methods that can accurately gauge the individual and

combined contributions of ligand, protein, and solvent entropy to the system's conformational entropy, and relate them to structural alterations[90–92,44,76].Computational methods for modeling ensemble-averaged experimental data in X-ray crystallography and CryoEM have the potential to report on how the overall entropy of the system changes. Our lab has developed qFit[29,93], which automatically creates multiconformer models of proteins and ligands from X-ray and cryoEM data. As discussed above, these models can be used to calculate crystallographic and potentially cryoEM order parameters, connecting conformational entropy to structural models. Other multiconformer approaches, such as Xtrapol8 and FLEXR[94,95,96], and X-ray-restrained MD ensemble methods[97,98,99], can also be used to identify the structural basis of changes in conformational heterogeneity. Additionally, new tools are coming online to compare and combine diffraction intensities between experiments to more sensitively detect differences in structure and conformational heterogeniety[100,101]. In CryoEM, recent years have seen significant advances in distinguishing unique large conformations within a single experiment [95102,103]. When comparing structures based on experimental data, we should think of conformational change as differences in modeled coordinates greater than propagated error in coordinate precision. With improved resolution and experimental and computational techniques, we can now explore much smaller conformational changes. These advances will complement the work emerging from NMR. Further, molecular dynamics simulations, especially those that can leverage experimental priors[25], will play an increasing role in illuminating connections between structure and conformational entropy.

Combining multiple experimental data sources will help elucidate the interplay between what are traditionally called “conformational changes” and redistributions of conformational ensembles. Along with these advancements, we need better data structures to capture ensemble based models to capture more accurate models and provide critical data for protein structural ensemble prediction or protein design. While we can now obtain highly accurate predictions of single conformer structure[104,105], these are poor representations of the conformational ensemble underlying proteins[106]. While algorithms are being developed to coax structure prediction software into modeling multiple states, it is unclear how best to implement or interpret these findings[107–109]. There are some promising approaches to integrate data from experiments and simulation which can help bridge this gap[110,111], especially MD reweighting procedures and AlphaFold ensembles[112,113]. By providing a baseline of training data better capturing the underlying ensembles we can move towards structural ensemble predictions[106], which will likely have functional implication. Using similar principles as protein structural prediction, protein design has found success in making stable structures. The next frontier is dynamics, which will enable design of even more innovative functions[114,115].

Now is the moment to start integrating the impact of conformational entropy into ligand and protein design. Traditional structure-based drug design emphasizes optimizing the complementarity of a small molecule into the ground state of a binding site. While this might

improve enthalpy, it almost completely ignores entropy (except for consideration of solvent entropy). But the spatial connection of conformational entropy of ligands is clear, even with tiny changes in chemical attributes leading to widely different entropies and in some cases, affinities. Using structural models that can combine structure and interactions of ligand with proteins along with the impact of conformational entropy will help usher in the use of this untapped energetic reservoir in drug design. Deciphering the foundations of conformational entropy's structural implications will pave the way for manipulating protein function, tracing protein evolution, and rationally engineering proteins and drugs.

CONCLUSION

In 1984, Cooper and Dryden proposed the hypothesis of allostery without conformational change[12]. Over four decades of extensive research necessitates a reevaluation of our understanding of 'conformational change'. Conformational changes vary in magnitude, ranging from subangstrom to many nanometer motions, spanning diverse timescales, encompassing conformational entropy differences. Emerging allosteric models demonstrate how the spectrum of conformational shifts collaboratively alters the conformational ensemble of a protein to regulate its function[8].

Within globular domains, our review highlights two main mechanisms to consider conformational entropy for design purposes: pre-ordering and spatial compensation. Pre-paying entropic costs by establishing order in the ground state has a long history, especially in enzymes. This mechanism is likely easier to engineer. One could imagine designing the pre-organized into the ground state. The second principle, spatial entropic compensation, which is achieved by increasing conformational entropy in distal regions following binding, is harder to conceptualize for design purposes. However, we now have the tools and framework to start chipping away at this problem. Finally, it is possible (and likely!) that there are additional dimensions by which conformational entropy can be considered in understanding and designing function. The body of research presented here is just starting to scratch the surface of the multiplicity of mechanisms that selection has used to drive free energy changes.

Acknowledgements

We thank Aashish Manglik for helpful comments on the GPCR section, Siyuan Du and Pei Zhou for helpful feedback. Our work is supported by NIH GM145238 (JSF) and U19AI171110 (JSF and SAW).

References

1. Makhatadze, G. I. & Privalov, P. L. On the entropy of protein folding. *Protein Sci.* **5**, 507–510 (1996).

2. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* vol. 7 95–99 Preprint at https://doi.org/10.1016/s0022-2836(63)80023-6 (1963).
3. Doig, A. J. & Sternberg, M. J. Side-chain conformational entropy in protein folding. *Protein Sci.* **4**, 2247–2251 (1995).
4. Tokuriki, N. & Tawfik, D. S. Protein dynamism and evolvability. *Science* **324**, 203–207 (2009).
5. Smith, J. L., Hendrickson, W. A., Honzatko, R. B. & Sheriff, S. Structural heterogeneity in protein crystals. *Biochemistry* **25**, 5018–5027 (1986).
6. Bromberg, S. & Dill, K. A. Side-chain entropy and packing in proteins. *Protein Science* vol. 3 997–1009 Preprint at https://doi.org/10.1002/pro.5560030702 (1994).
7. Shapovalov, M. V. & Dunbrack, R. L., Jr. A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* **19**, 844–858 (2011).
8. Motlagh, H. N., Wrabl, J. O., Li, J. & Hilser, V. J. The ensemble nature of allostery. *Nature* **508**, 331–339 (2014).
9. Wankowicz, S. A., de Oliveira, S. H., Hogan, D. W., van den Bedem, H. & Fraser, J. S. Ligand binding remodels protein side-chain conformational heterogeneity. *Elife* **11**, (2022).
10. Hilser, V. J., Dowdy, D., Oas, T. G. & Freire, E. The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9903–9908 (1998).
11. Monod, J., Wyman, J. & Changeux, J. P. ON THE NATURE OF ALLOSTERIC TRANSITIONS: A PLAUSIBLE MODEL. *J. Mol. Biol.* **12**, 88–118 (1965).

12. Cooper, A. & Dryden, D. T. Allostery without conformational change. A plausible model. *Eur. Biophys. J.* **11**, 103–109 (1984).
13. Wand, A. J. The dark energy of proteins comes to light: conformational entropy and its role in protein function revealed by NMR relaxation. *Curr. Opin. Struct. Biol.* **23**, 75–81 (2013).
14. Wand, A. J. & Sharp, K. A. Measuring Entropy in Molecular Recognition by Proteins. *Annu. Rev. Biophys.* **47**, 41–61 (2018).
15. Kasinath, V., Sharp, K. A. & Wand, A. J. Microscopic insights into the NMR relaxation-based protein conformational entropy meter. *J. Am. Chem. Soc.* **135**, 15092–15100 (2013).
16. Igumenova, T. I., Frederick, K. K. & Wand, A. J. Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution. *Chem. Rev.* **106**, 1672–1699 (2006).
17. [No title]. https://doi.org/10.1021/ja00381a009.
18. Lipari, G. & Szabo, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *Journal of the American Chemical Society* vol. 104 4559–4570 Preprint at https://doi.org/10.1021/ja00381a010 (1982).
19. Frederick, K. K., Marlow, M. S., Valentine, K. G. & Wand, A. J. Conformational entropy in molecular recognition by proteins. *Nature* **448**, 325–329 (2007).
20. Caro, J. A. *et al.* Entropy in molecular recognition by proteins. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 6563–6568 (2017).
21. Marques, B. S. *et al.* Protein conformational entropy is not slaved to water. *Sci. Rep.* **10**, 1–8 (2020).
22. Hoffmann, F., Mulder, F. A. A. & Schäfer, L. V. Predicting NMR relaxation of proteins from

molecular dynamics simulations with accurate methyl rotation barriers. *J. Chem. Phys.* **152**, 084102 (2020).

23. McDonald, L. R., Whitley, M. J., Boyer, J. A. & Lee, A. L. Colocalization of fast and slow timescale dynamics in the allosteric signaling protein CheY. *J. Mol. Biol.* **425**, 2372–2381 (2013).
24. Ray, K. K. *et al.* Entropic control of the free-energy landscape of an archetypal biomolecular machine. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2220591120 (2023).
25. Bottaro, S. & Lindorff-Larsen, K. Biophysical experiments and biomolecular simulations: A perfect match? *Science* **361**, 355–360 (2018).
26. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).
27. Glaeser, R. M. How Good Can Single-Particle Cryo-EM Become? What Remains Before It Approaches Its Physical Limits? *Annu. Rev. Biophys.* **48**, 45–61 (2019).
28. Keedy, D. A. *et al.* An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *Elife* **7**, (2018).
29. Wankowicz, S. A. *et al.* Uncovering Protein Ensembles: Automated Multiconformer Model Building for X-ray Crystallography and Cryo-EM. *bioRxiv* (2023) doi:10.1101/2023.06.28.546963.
30. Fenwick, R. B., van den Bedem, H., Fraser, J. S. & Wright, P. E. Integrated description of protein dynamics from room-temperature X-ray crystallography and NMR. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E445–54 (2014).
31. Lee, A. L., Kinnear, S. A. & Wand, A. J. Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex. *Nat. Struct. Biol.* **7**, 72–77 (2000).

32. Popovych, N., Sun, S., Ebright, R. H. & Kalodimos, C. G. Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* **13**, 831–838 (2006).
33. Tzeng, S.-R. & Kalodimos, C. G. Protein activity regulation by conformational entropy. *Nature* **488**, 236–240 (2012).
34. Tzeng, S.-R. & Kalodimos, C. G. Dynamic activation of an allosteric regulatory protein. *Nature* **462**, 368–372 (2009).
35. Caro, J. A., Valentine, K. G., Cole, T. R. & Wand, A. J. Pressure, motion, and conformational entropy in molecular recognition by proteins. *Biophys Rep (N Y)* **3**, 100098 (2023).
36. Whitten, S. T., García-Moreno E, B. & Hilser, V. J. Local conformational fluctuations can modulate the coupling between proton binding and global structural transitions in proteins. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 4282–4287 (2005).
37. Petit, C. M., Zhang, J., Sapienza, P. J., Fuentes, E. J. & Lee, A. L. Hidden dynamic allostery in a PDZ domain. *Proceedings of the National Academy of Sciences* vol. 106 18249–18254 Preprint at https://doi.org/10.1073/pnas.0904492106 (2009).
38. Lee, H.-J. & Zheng, J. J. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun. Signal.* **8**, 8 (2010).
39. Cui, Q. & Karplus, M. Allostery and cooperativity revisited. *Protein Sci.* **17**, 1295–1307 (2008).
40. Capdevila, D. A., Braymer, J. J., Edmonds, K. A., Wu, H. & Giedroc, D. P. Entropy redistribution controls allostery in a metalloregulatory protein. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 4424–4429 (2017).
41. Bonin, J. P. *et al.* Positive Cooperativity in Substrate Binding by Human Thymidylate

Synthase. *Biophys. J.* **120**, 4137 (2021).

42. Bonin, J. P., Sapienza, P. J. & Lee, A. L. Dynamic allostery in substrate binding by human thymidylate synthase. *Elife* **11**, (2022).
43. Smith, C. A. *et al.* Population shuffling of protein conformations. *Angew. Chem. Int. Ed Engl.* **54**, 207–210 (2015).
44. Kim, T. H. *et al.* The role of dimer asymmetry and protomer dynamics in enzyme catalysis. *Science* **355**, (2017).
45. Sapienza, P. J., Falk, B. T. & Lee, A. L. Bacterial Thymidylate Synthase Binds Two Molecules of Substrate and Cofactor without Cooperativity. *J. Am. Chem. Soc.* **137**, 14260–14263 (2015).
46. Olivieri, C. *et al.* ATP-competitive inhibitors modulate the substrate binding cooperativity of a kinase by altering its conformational entropy. *Sci Adv* **8**, eabo0696 (2022).
47. Pucheta-Martínez, E. *et al.* An Allosteric Cross-Talk Between the Activation Loop and the ATP Binding Site Regulates the Activation of Src Kinase. *Sci. Rep.* **6**, 24235 (2016).
48. Xie, T., Saleh, T., Rossi, P. & Kalodimos, C. G. Conformational states dynamically populated by a kinase determine its function. *Science* **370**, (2020).
49. Kim, C. *et al.* A biophysical framework for double-drugging kinases. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2304611120 (2023).
50. Beyett, T. S. *et al.* Molecular basis for cooperative binding and synergy of ATP-site and allosteric EGFR inhibitors. *Nat. Commun.* **13**, 2530 (2022).
51. Ben Haddou, T. *et al.* Pharmacological investigations of N-substituent variation in morphine and oxymorphone: opioid receptor binding, signaling and antinociceptive activity. *PLoS One* **9**, e99231 (2014).

52. Powers, A. S. *et al.* Structural basis of efficacy-driven ligand selectivity at GPCRs. *Nat. Chem. Biol.* **19**, 805–814 (2023).
53. Nygaard, R. *et al.* The dynamic process of β(2)-adrenergic receptor activation. *Cell* **152**, 532–542 (2013).
54. Clark, L. D. *et al.* Ligand modulation of sidechain dynamics in a wild-type human GPCR. *Elife* **6**, (2017).
55. Bumbak, F. *et al.* Ligands selectively tune the local and global motions of neurotensin receptor 1 (NTS). *Cell Rep.* **42**, 112015 (2023).
56. O'Brien, E. S. *et al.* Membrane Proteins Have Distinct Fast Internal Motion and Residual Conformational Entropy. *Angew. Chem. Int. Ed Engl.* **59**, 11108–11114 (2020).
57. Pontiggia, F. *et al.* Free energy landscape of activation in a signalling protein at atomic resolution. *Nat. Commun.* **6**, 7284 (2015).
58. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
59. Sahni, N. *et al.* Widespread macromolecular interaction perturbations in human genetic disorders. *Cell* **161**, 647–660 (2015).
60. Armenia, J. *et al.* The long tail of oncogenic drivers in prostate cancer. *Nat. Genet.* **50**, 645–651 (2018).
61. Bigman, L. S. & Levy, Y. Entropic contributions to protein stability. *Isr. J. Chem.* **60**, 705–712 (2020).
62. Hamborg, L. *et al.* Synergistic stabilization of a double mutant in chymotrypsin inhibitor 2 from a library screen in E. coli. *Commun Biol* **4**, 980 (2021).
63. Gavrilov, Y. *et al.* Double Mutant of Chymotrypsin Inhibitor 2 Stabilized through Increased

Conformational Entropy. *Biochemistry* **61**, 160–170 (2022).

64. Saavedra, H. G., Wrabl, J. O., Anderson, J. A., Li, J. & Hilser, V. J. Dynamic allostery can drive cold adaptation in enzymes. *Nature* **558**, 324–328 (2018).
65. Jiang, F. *et al.* Creatinase: Using Increased Entropy to Improve the Activity and Thermostability. *J. Phys. Chem. B* **127**, 2671–2682 (2023).
66. Liu, Y. *et al.* The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. *Nature* **602**, 294–299 (2022).
67. Prévost, J. *et al.* Impact of temperature on the affinity of SARS-CoV-2 Spike glycoprotein for host ACE2. *J. Biol. Chem.* **297**, 101151 (2021).
68. Vergara, N. G., Gatchel, M. & Abrams, C. F. Entropic overcompensation of the N501Y mutation on SARS-CoV-2 S binding to ACE2. *bioRxiv* (2022) doi:10.1101/2022.08.30.505841.
69. Mikolajek, H. *et al.* Correlation between the binding affinity and the conformational entropy of nanobody SARS-CoV-2 spike protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2205412119 (2022).
70. Yabukarski, F. *et al.* Assessment of enzyme active site positioning and tests of catalytic mechanisms through X-ray-derived conformational ensembles. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 33204–33215 (2020).
71. Yabukarski, F. *et al.* Ensemble-function relationships to dissect mechanisms of enzyme catalysis. *Sci Adv* **8**, eabn7738 (2022).
72. Korostelev, A. A. The Structural Dynamics of Translation. *Annu. Rev. Biochem.* **91**, 245–267 (2022).
73. Sievers, A., Beringer, M., Rodnina, M. V. & Wolfenden, R. The ribosome as an entropy trap.

Proc. Natl. Acad. Sci. U. S. A. **101**, 7897–7901 (2004).

74. Kaledhonkar, S. *et al.* Late steps in bacterial translation initiation visualized using time-resolved cryo-EM. *Nature* **570**, 400–404 (2019).
75. Keul, N. D. *et al.* The entropic force generated by intrinsically disordered segments tunes protein function. *Nature* **563**, 584–588 (2018).
76. Capdevila, D. A. *et al.* Functional Role of Solvent Entropy and Conformational Entropy of Metal Binding in a Dynamically Driven Allosteric System. *J. Am. Chem. Soc.* **140**, 9108–9119 (2018).
77. Gorman, S. D., Winston, D. S., Sahu, D. & Boehr, D. D. Different Solvent and Conformational Entropy Contributions to the Allosteric Activation and Inhibition Mechanisms of Yeast Chorismate Mutase. *Biochemistry* **59**, 2528–2540 (2020).
78. Yu, F. & Sukenik, S. Structural preferences shape the entropic force of disordered protein ensembles. *bioRxiv* (2023) doi:10.1101/2023.01.20.524980.
79. Borgia, A. *et al.* Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61–66 (2018).
80. Skriver, K., Theisen, F. F. & Kragelund, B. B. Conformational entropy in molecular recognition of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **83**, 102697 (2023).
81. White, J. T., Li, J., Grasso, E., Wrabl, J. O. & Hilser, V. J. Ensemble allosteric model: energetic frustration within the intrinsically disordered glucocorticoid receptor. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **373**, (2018).
82. Skaist Mehlman, T. *et al.* Room-temperature crystallography reveals altered binding of small-molecule fragments to PTP1B. *Elife* **12**, (2023).
83. van Zundert, G. C. P. *et al.* qFit-ligand Reveals Widespread Conformational Heterogeneity

of Drug-Like Molecules in X-Ray Electron Density Maps. *J. Med. Chem.* **61**, 11183–11198 (2018).

84. Mobley, D. L., Chodera, J. D. & Dill, K. A. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* **125**, 084902 (2006).
85. Lee, C.-J. *et al.* Drug design from the cryptic inhibitor envelope. *Nat. Commun.* **7**, 10638 (2016).
86. Zhou, P. & Hong, J. Structure- and Ligand-Dynamics-Based Design of Novel Antibiotics Targeting Lipid A Enzymes LpxC and LpxH in Gram-Negative Bacteria. *Acc. Chem. Res.* **54**, 1623–1634 (2021).
87. Abel, R., Young, T., Farid, R., Berne, B. J. & Friesner, R. A. Role of the active-site solvent in the thermodynamics of factor Xa ligand binding. *J. Am. Chem. Soc.* **130**, 2817–2831 (2008).
88. Hilser, V. J., García-Moreno E, B., Oas, T. G., Kapp, G. & Whitten, S. T. A statistical thermodynamic model of the protein ensemble. *Chem. Rev.* **106**, 1545–1558 (2006).
89. Majewski, M., Ruiz-Carmona, S. & Barril, X. An investigation of structural stability in protein-ligand complexes reveals the balance between order and disorder. *Communications Chemistry* **2**, 1–8 (2019).
90. Bruning, J. B. *et al.* Coupling of receptor conformation and ligand orientation determine graded activity. *Nat. Chem. Biol.* **6**, 837–843 (2010).
91. Srinivasan, S. *et al.* Ligand-binding dynamics rewire cellular signaling via estrogen receptor-α. *Nat. Chem. Biol.* **9**, 326–332 (2013).
92. Qu, Q. *et al.* Insights into distinct signaling profiles of the μOR activated by diverse

agonists. *Nat. Chem. Biol.* **19**, 423–430 (2023).

93. Riley, B. T. *et al.* qFit 3: Protein and ligand multiconformer modeling for X-ray crystallographic and single-particle cryo-EM density maps. *Protein Sci.* **30**, 270–285 (2021).
94. De Zitter, E., Coquelle, N., Oeser, P., Barends, T. R. M. & Colletier, J.-P. Xtrapol8 enables automatic elucidation of low-occupancy intermediate-states in crystallographic studies. *Commun Biol* **5**, 640 (2022).
95. Stachowski, T. R. & Fischer, M. FLEXR: automated multi-conformer model building using electron-density map sampling. *Acta Crystallogr D Struct Biol* **79**, 354–367 (2023).
96. Lang, P. T. *et al.* Automated electron-density sampling reveals widespread conformational polymorphism in proteins. *Protein Sci.* **19**, 1420–1431 (2010).
97. Burnley, B. T., Tom Burnley, B., Afonine, P. V., Adams, P. D. & Gros, P. Author response: Modelling dynamics in protein crystal structures by ensemble refinement. Preprint at https://doi.org/10.7554/elife.00311.029 (2012).
98. Pearce, N. M. & Gros, P. A method for intuitively extracting macromolecular dynamics from structural disorder. *Nat. Commun.* **12**, 5493 (2021).
99. Ploscariu, N., Burnley, T., Gros, P. & Pearce, N. M. Improving sampling of crystallographic disorder in ensemble refinement. *Acta Crystallogr D Struct Biol* **77**, 1357–1364 (2021).
100. Dalton, K. M., Greisman, J. B. & Hekstra, D. R. A unifying Bayesian framework for merging X-ray diffraction data. *Nat. Commun.* **13**, 7764 (2022).
101. Pearce, N. M. *et al.* A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat. Commun.* **8**, 15123 (2017).
102. Levy, A., Wetzstein, G., Martel, J., Poitevin, F. & Zhong, E. D. Amortized Inference for Heterogeneous Reconstruction in Cryo-EM. (2022) doi:10.48550/arXiv.2210.07387.

103. Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).
104. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
105. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
106. Lane, T. J. Protein structure prediction has reached the single-structure frontier. *Nat. Methods* (2023) doi:10.1038/s41592-022-01760-4.
107. Ingraham, J. *et al.* Illuminating protein space with a programmable generative model. *bioRxiv* 2022.12.01.518682 (2022) doi:10.1101/2022.12.01.518682.
108. Del Alamo, D., Sala, D., Mchaourab, H. S. & Meiler, J. Sampling alternative conformational states of transporters and receptors with AlphaFold2. *Elife* **11**, (2022).
109. Vani, B. P., Aranganathan, A., Wang, D. & Tiwary, P. AlphaFold2-RAVE: From sequence to Boltzmann ensemble. *bioRxiv* 2022.05.25.493365 (2022) doi:10.1101/2022.05.25.493365.
110. Lindorff-Larsen, K., Best, R. B., Depristo, M. A., Dobson, C. M. & Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **433**, 128–132 (2005).
111. Olsson, S., Wu, H., Paul, F., Clementi, C. & Noé, F. Combining experimental and simulation data of molecular processes via augmented Markov models. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 8265–8270 (2017).
112. da Silva, G. M., Cui, J. Y., Dalgarno, D. C., Lisi, G. P. & Rubenstein, B. M. Predicting Relative Populations of Protein Conformations without a Physics Engine Using AlphaFold2. *bioRxiv* (2023) doi:10.1101/2023.07.25.550545.

113. Vani, B. P., Aranganathan, A., Wang, D. & Tiwary, P. AlphaFold2-RAVE: From Sequence to Boltzmann Ranking. *J. Chem. Theory Comput.* **19**, 4351–4354 (2023).

114. Kuhlman, B. & Bradley, P. Advances in protein structure prediction and design. *Nature Reviews Molecular Cell Biology* vol. 20 681–697 Preprint at https://doi.org/10.1038/s41580-019-0163-x (2019).

115. Choi, J. H., Laurent, A. H., Hilser, V. J. & Ostermeier, M. Design of protein switches based on an ensemble model of allostery. *Nat. Commun.* **6**, 6968 (2015).