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.
* mullane.stephanie@gmail.com, jfraser@fraserlab.com

ABSTRACT

During protein folding, proteins transition from a disordered polymer into a globular structure, markedly decreasing their conformational degrees of freedom and consequently leading to a substantial reduction in entropy. Nonetheless, folded proteins still retain significant entropy as they fluctuate between the conformations that make up their native state. This residual entropy contributes to crucial functions like binding or catalysis. Here, we outline three major ways that protein use conformational entropy to perform their functions; first, pre-paying entropic cost through ordering of the ground state; second, spatially redistributing entropy, where an decrease in entropy in one area is reciprocated by an increase in entropy elsewhere; third, populating catalytically-competent ensembles, where conformational entropy within the enzymatic scaffold aids in lowering transition state barriers. Given the growing evidence of the biological significance of conformational entropy, emerging largely from NMR and simulation studies, solving the current challenge of structurally defining the ensembles encoding conformational entropy will open new paths for control of binding, catalysis, and allostery.

Proteins retain significant conformational entropy within their folded state

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 conformations, especially in backbone atoms, is restricted by steric clashes[2]. The hydrophobic collapse accompanying folding also liberates water molecules from ordering around side chains. This collapse is entropically favorable as the water molecules acquire additional degrees of freedom[3]. However, the substantial reduction in conformational freedom of the protein that accompanies folding has a steep entropic cost. This entropic loss is governed by the Boltzmann entropy equation ([$S=k_B\ln(W)$]; S=entropy; k_B=Boltzmann's constant; W=multiplicity of accessible conformations)[4]. Additional free energy favoring folding is due to attractive interactions (e.g. hydrogen bonds), which must outweigh the cost of further reducing the number of accessible conformations that forming those interactions requires (**Figure 1**).

The remaining conformations folded proteins access provides residual entropy, required for protein stability and activity. A strategy for maximizing the stability of the folded state is therefore to balance

the attractive interactions (enthalpy) with as minimal a reduction in accessible conformations (entropy) as possible. In the conformational ensemble of the folded native state for most globular proteins, conformational heterogeneity is usually dominated by alternative conformations of side chains, with only minor influences from the backbone[5]. Perturbations, such as ligand binding, change the conformational ensemble and can alter the residual conformational entropy. The diversity of the conformational ensemble also correlates with the protein's potential for evolutionary adaptation[6].

To concretely illustrate how the residual conformational entropy in the native state can influence the free energy of the system upon perturbation, consider a single leucine side chain near the active site of a protein kinase (**Figure 1**). In the unfolded state, the leucine side chain can equally access all rotamers[7], and the backbone is only restricted by the neighboring residues on the chain (**Figure 1A**). Upon folding, the backbone becomes heavily restricted by steric clashes with other residues and adopts a relatively unique conformation[2]. This loss of backbone freedom incurs a significant entropic penalty, with residual entropy contributed only by the harmonic motion about its mean position. Unlike the backbone, side chains can be much less restricted, and most side chains access multiple rotameric states, even within the protein core[8]. Without a ligand in the active site, the leucine side chain can access two of the five possible rotameric states, while steric clashes with other residues disfavor the remaining one. This scenario leads to a reduction, but not the elimination, of side chain conformational entropy upon folding (**Figure 1B**).

Perturbations, including ligand binding, can further alter the conformational ensemble with varying entropic consequences[9]. An active site ligand may form van der Waals contacts with the leucine side chain in a way that strongly prefers one of the apo state alternative conformations (**Figure 1C**). The transition from two to one rotameric conformation further decreases conformational entropy. Since most structural models from X-ray crystallography or cryoEM would only have one of the two conformations modeled in the apo state, whether the ligand binding induces a “conformational change” depends on which conformation is modeled in the apo state (usually, but not always, the more populated one). But, by considering the underlying ensemble, we can see that ligand binding induces a “conformational redistribution” and can start to think about the entropic consequences on the free energy of binding.

Ligand binding can also, counterintuitively, increase the conformational entropy of side chains[10,11]. For example, a ligand binding at a different part of the active site may stabilize different residues in conformations that create a small void adjacent to the leucine. The void could permit populating an additional, third, leucine rotameric conformation compared to apo (**Figure 1D**) or increase the thermal fluctuations around the two already populated rotameric conformations (**Figure 1E**). Both of these scenarios increase conformational entropy and can, therefore, favor ligand binding.

While the entropic contribution of a single side chain may appear small, accumulating numerous conformational redistributions across the protein can lead to a significant entropic impact and provide an allosteric source of free energy. Structural descriptions of allostery are frequently based on averaged single conformer structures and described in a “domino effect” of conformational changes through a well-defined pathway[12]. But 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[13]. This model can be reconciled with the domino effect by examining allostery through a thermodynamic lens where any change in the conformational ensemble, whether large or small, can impact free energy and protein function through both entropy and enthalpy[9].

Despite the apparent energetic importance of conformational entropy, the ability to probe it experimentally is underdeveloped. Currently, the best experimental support comes from NMR relaxation techniques. By probing methyl or amide groups, NMR relaxation techniques obtain a site-specific measurement of the amount of disorder on the pico-nanosecond timescale, called an order parameter[14]. Changes in the average side chain order parameter for methyl-containing residues (like leucine) across a protein have been correlated with calorimetric measurements of conformational entropy[14–18]. Excitingly, these experiments reveal a great diversity of behaviors across systems: some binding events are entropically driven due to increased side chain disorder whereas others are strongly entropically disfavored. Additionally, some side chains exhibit increased disorder even when the overall trend is towards unfavorable conformational entropy (and vice versa). Furthermore, different ligands bound to the same receptor can have drastically different impacts on conformational entropy[10,19,20]. While NMR order parameters provide the extent of the disorder, many equivalent side chain conformational ensembles can produce a given order parameter value, as seen in the packing void example (**Figure 1D,E**)[17]. This degeneracy makes it challenging to connect conformational entropy to structural models.

An alternative, and readily available, source to reveal the structural basis of conformational entropy is 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 exchange timescale. Heterogeneity of protein side chains, which contain the majority of protein conformational entropy, is often not disruptive to diffraction or alignment. While careful attention to experimental noise is needed[21,22], we have demonstrated that multiconformer modeling in high-resolution X-ray datasets can detect subtle and weak conformations, explaining statistically significant signals in the density map[23]. The leucine side chain would have subtly different density surrounding it in each scenario. While a multiconformer model would help tease apart the contributions of different rotamers, a conventional single-state structure would almost certainly model all four examples the same way. A "crystallographic order parameter" that agrees reasonably with NMR relaxation experiments can be obtained by determining the distance between each alternative conformation and the weighted B-factor, a refined parameter to account for harmonic disorder, of each conformer[24]. Importantly, it is not limited to methyl-containing side chains and, due to the time-averaged nature of diffraction or single particle imaging, this metric extends to timescales beyond the pico-nanosecond timescale of NMR order parameters. Ideally, molecular dynamics (MD) simulations could directly observe the fluctuations of residue upon perturbation[25], but significant protein dynamics happen on a timescale inaccessible to MD[26–28], necessitating hybrid approaches that bias agreement to experimentally observed order parameters[29]. Collectively, these techniques are beginning to quantify thstructural basis of conformational entropy.

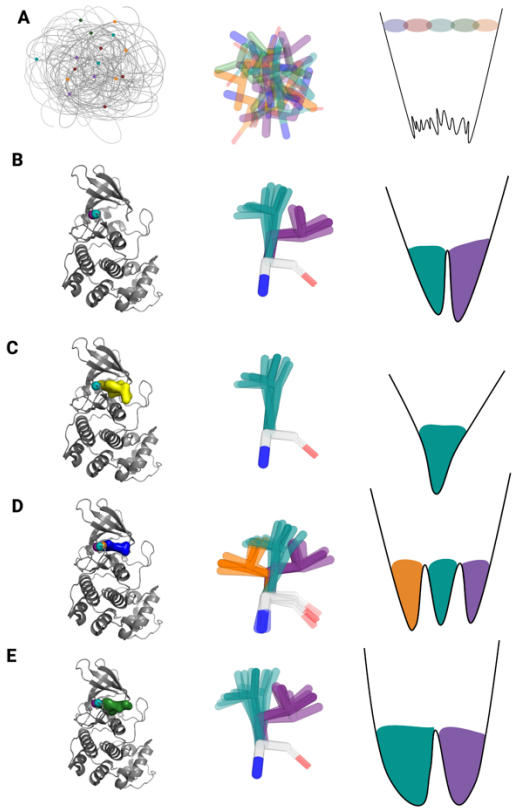

Figure 1. Examples of an estimated conformational entropy for a leucine side chain in different contexts. (Left) The structural state of the entire protein with the leucine showed in spheres. (Middle) The conformation(s) of the leucine. (Right) The relative location of each conformation on the energy landscape of the leucine side chain. The entropic contribution of the leucine is characterized by the Boltzmann entropy equation ([$S=k_B\ln(W)$]; S=entropy; k_B=Boltzmann's constant; W=number of conformations)[4] and additional contributions from the width of the basins under a harmonic approximation[30].
(A) In an unfolded state, the entire protein and the leucine side chain access many different conformations, with each conformation sitting at the top of a wide energy landscape for the entire protein.
(B) Upon folding, the majority of the protein is structured. This results in a large loss of entropy in the leucine side chain. The leucine can only access two states, separated by an energy barrier. One conformation has significantly more harmonic entropy, as shown by a wider energy well (teal) relative to the other conformation (purple). B-factors model this harmonic motion in X-ray crystallography. The simplified energy landscape of the leucine side chain alone is shown.
(C) When one ligand binds, the leucine 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 with the leucine residue, the leucine gains conformational entropy by accessing three of five 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. This would observed as an increased disorder in NMR order parameters, but 3 alternative conformations in X-ray or cryoEM data.
(E) Another example of a leucine gaining conformational entropy upon ligand binding. While the leucine has the same number of accessible rotamer states as in the apo state (2), the harmonic motion in both rotamer states greatly increases. This would be observed as an increase in disorder in NMR order parameters, but 2 alternative conformations with increased B-factors in X-ray or cryoEM data.

Our prevailing models for visualizing proteins tend to focus on explicitly representing and considering enthalpic interactions, emphasizing the precise orientation of "functional" residues. However, entropy is often relegated to a more implicit role in the analysis. We need to consider how a local narrowing of the conformational ensemble is opposed by entropy. This opens the question, what strategies do proteins employ to optimize the overall free energy by optimizing the entropic contributions?

Here, we review evidence of how nature has evolved to harness protein conformational entropy. The first principle is most well known: *pre-paying entropic costs* through ordering in the initial, or ground, state. The second phenomenon is *spatial entropic redistribution*: increasing conformational entropy in one region of the protein or system in response to a decrease in conformational entropy elsewhere in the system. The third concept extends the concept of spatial redistribution to *catalytically competent ensembles*: conformational entropy in the enzymatic scaffold, distant from the catalytic site, contributes to lowering transition state barriers. Further explorations of these mechanisms could profoundly impact our understanding of binding events, catalysis, and cell signaling. After all, without explicitly considering conformational entropy, we ignore a large thermodynamic component driving free energy. Now is the time to apply a conformational entropy lens to our explanation of protein function and begin applying that lens prospectively to ligand and protein design.

Pre-paying entropic penalties

Binding affinity is largely dictated by the surface complementarity and attractive interactions (e.g. hydrogen bonds) between the ligand and the receptor. Like folding, both these features greatly constrain the potential conformations the receptor can adopt, incurring an entropic penalty. The loss of conformational entropy is countered largely by the free energy gained by these interactions. However, if the ground state conformational ensemble is restricted to a similar extent to those adopted in the bound state, the protein *pre-pays the entropic penalty*. The cost of the entropic penalty can be minimized or even rendered non-existent. This simple concept extends most interestingly to cooperativity in allostery.

Consider an allosteric system with two binding sites, each of which must be well-ordered to stably bind a ligand. The first binding event significantly reduces the conformational freedom of side chains across the protein, paying a large entropic penalty[31]. This reduction in conformational heterogeneity can stabilize binding-competent conformations at the second binding site. In this scenario, the first event ‘pre-pays’ the entropic penalty of the second binding event and reduces the entropic organization penalty for stabilizing the binding-competent conformation. Compared to a scenario where the second binding site remains dynamic, allosteric ordering lowers the energy gap for the second binding event, leading to an observation of ‘positive cooperativity’(**Figure 2A**).

A beautiful example of how entropy influences positive cooperativity is in the dimeric enzyme human thymidylate synthase (hTS)[19,32]. Careful calorimetric measurements at multiple protein concentrations demonstrated positive cooperativity between sites, with the second binding event requiring ~1.5kcal/mol less energy than the first. This difference is driven primarily by entropy. Both crystallography and sensitive NMR measurements revealed no protein structural changes upon ligand binding. However, NMR measurements of side chain order parameters suggest that the first binding event greatly reduces conformational entropy throughout the protein, including at the second binding site. This ordering leads to virtually no change in protein conformational entropy after the second binding event, which is consistent with calorimetry indicating that the second binding event is more favored entropically (**Figure 2B**). In addition, NMR relaxation dispersion and Chemical Exchange Saturation Transfer (CEST) experiments revealed a reduction in the exchange of some lowly populated states (~1%) upon ligand saturation at both sites. However, this population change is insufficient to explain the large swing in entropy. The loss of slow dynamics and dramatic change in side chain dynamics can be reconciled through population shuffling, where the fast dynamics undergo

reorganization due to the influence of slower movements[19,33]. Remarkably, the entropy-driven positive cooperativity is highly specific to hTS's substrate dUMP. Even though the substrate (dUMP) and product (dTMP) differ by only a single methyl group, no positive cooperativity is observed by calorimetry for dTMP binding. Again, the NMR data are consistent with this observation, showing no difference in side chain order parameters between hTS-TMP and apo hTS. These observations link the thermodynamics of cooperativity to the pre-ordering of side chain conformations across the protein.

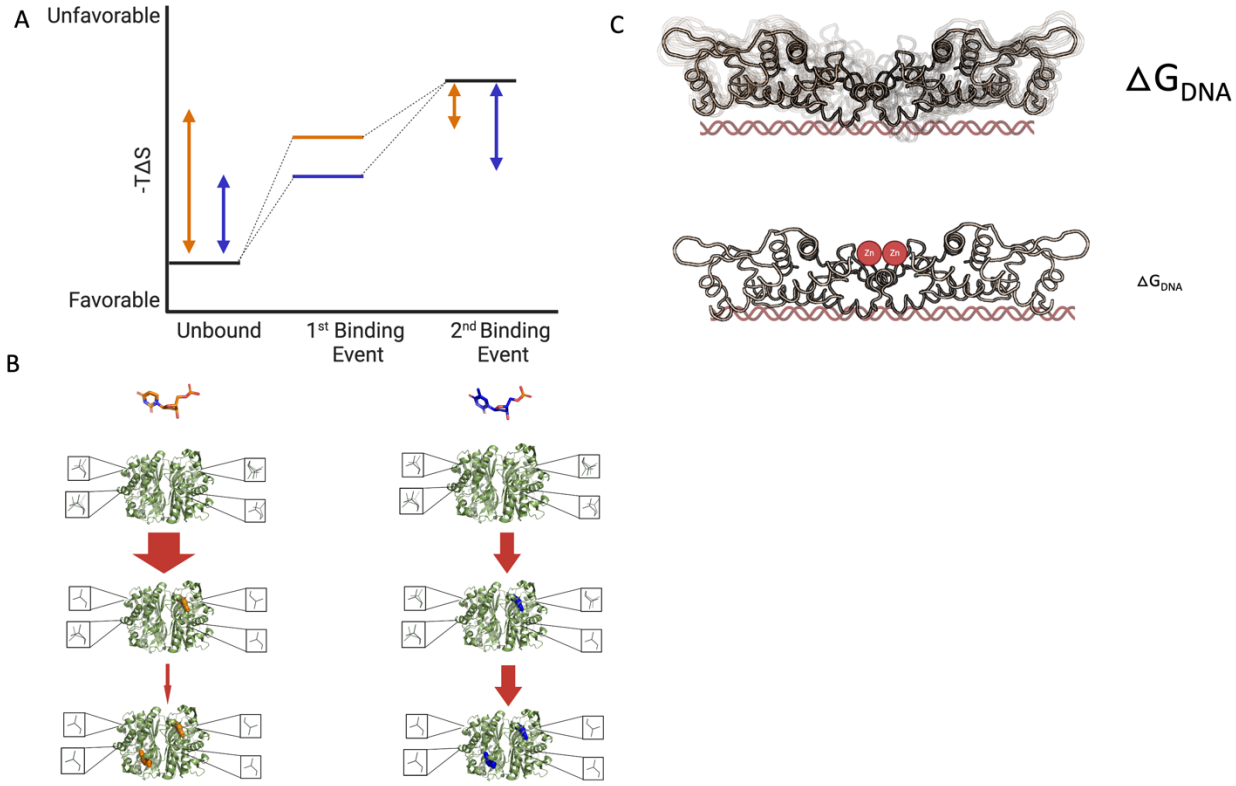

Figure 2. *(A) Example entropy changes over a cooperativity event. In entropically driven positive cooperativity (orange), a significant entropic penalty occurs upon the initial binding event, evidenced by a substantial reduction in entropy (likely leading to some pre-ordering in the second binding site). This allows the second binding event to have a much smaller reduction in entropy, increasing the relative binding affinity and leading to positive cooperativity. In contrast (blue), the initial binding event only pays the entropic penalty of the first binding event, leading to the second binding event paying the same entropic penalty, preventing entropically driven positive cooperativity.*

(B) Human thymidylate synthase (hTS) displays entropically driven positive cooperativity when binding its substrate dUMP (orange). 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). When hTS binds to its product dTMP (blue), no positive cooperativity is observed (both binding events have similar binding affinities), with the entropy decreasing upon both binding events.

(C) In the bacterial transcription factor CzrA, without zinc, increased side chain flexibility in CrzA's core enhances DNA binding affinity due to favorable entropy. However, zinc binding reduces this flexibility, diminishing the favorable entropy and lowering DNA binding affinity.

Entropy can also contribute to cooperativity in monomeric proteins with multiple binding sites, as observed in protein kinases[34–36]. For example, in protein kinase A, cooperativity is observed between the nucleotide and substrate (peptide) binding sites[34]. As in hTS, this cooperativity has an entropic origin with pre-paying entropic costs and pre-organization as dominant molecular mechanisms.

Inhibitors that resemble the natural substrate, ATP, lead to ordering at the substrate binding site and exhibit greater apparent cooperativity[34]. This example highlights how evolution (and, perhaps unwittingly, medicinal chemistry efforts) have manipulated conformational entropy to achieve positive cooperativity.

In contrast to the positive cooperativity discussed above, entropy can also drive negative cooperativity. In the homotetramer transthyretin (TTR), the affinity for small molecules to the second site is weaker. However, the structure of each subunit, including the binding sites, looks structurally similar, posing the question of why there is a difference in affinity. CryoEM revealed an inherent asymmetry in the ensemble of conformations of each dimer and that the first binding event does not lead to ordering the second binding site. On the contrary, the subsequent binding event requires more extensive structural rearrangements, including loop ordering, which comes with at a higher entropic cost[37]. Entropically driven negative cooperativity is also observed with NMR order parameter measurements in the bacterial transcription factor CzrA, which has a Zinc binding site more than 15 angstroms away from a DNA-binding site. Without Zinc, DNA binding increases side chain flexibility throughout the core of CrzA, increasing DNA binding affinity through favorable entropy. However, upon binding Zinc, side chain flexibility strongly decreases, and thus conformational entropy, eliminating this favorable entropy and thus reducing DNA binding affinity (**Figure 2C**)[38].

These examples show how both negative and positive cooperativity can result from changes in the conformational distributions that accompany initial binding events. These variations in conformational distributions arise not only from significant shifts in their dominant conformation (conformational change) but also, crucially, from fluctuations. Pre-paying (or, in the case of negative cooperativity, exacerbating) entropic penalties is a key way proteins allosterically control function. A greater structural understanding of the entropic and reorganization penalties accompanying these binding events may be especially therapeutically relevant to double-drugging strategies[39,40].

Spatial entropic redistribution

While there are examples of “fuzzy” ligand-protein binding events where both the protein and ligand adopt many conformations upon binding[41], most binding events involve a local reduction in conformational dynamics at the binding site[10]. As in folding, this creates an entropic cost that the new attractive interactions between ligand and receptor must counter. Pre-paying this entropic cost with a ground state with a relatively less dynamic binding site is an attractive strategy, as stated above. However, such well-packed binding sites are at odds with the fact that most binding sites have large voids that can be filled by alternative conformations of binding site residues in the unbound form.

Although binding often comes with a local entropic cost, not all binding events have large overall entropic penalties. A significant contributor to counter the cost is surely the liberation of ordered water molecules in the unbound state; however, even after accounting for solvation changes, the contributions of protein conformational entropy can even favor binding[42]. This finding raises the hypothesis that nature counters the entropic penalty incurred at the binding site by modulating entropy elsewhere in the protein. Substantial evidence suggests that proteins have evolved to redistribute entropy from the binding site upon ligand interaction. This phenomenon, which we term *spatial entropic redistribution*, involves countering local ordering, manifested as a loss of entropy at the binding site, with an increase in disorder at more distant sites, reducing the overall change in

conformational entropy. Importantly, this is not a “zero sum” compensation, as in classic entropy-enthalpy compensation[43], but rather spatial entropy redistribution is an important mechanism for reducing the impact of unfavorable local entropic changes and, therefore, improving the overall free energy of binding. While pre-paying entropic penalties result from ordering a specific binding site in the ground state, spatial entropic redistribution results from distal sites increasing entropy in response to a local reduction at the binding site. The change in entropy from distal sites can, therefore, partially contribute to overcoming the entropy reduction in the binding site or even exceeding it, leading to a favorable change in conformational entropy of binding.

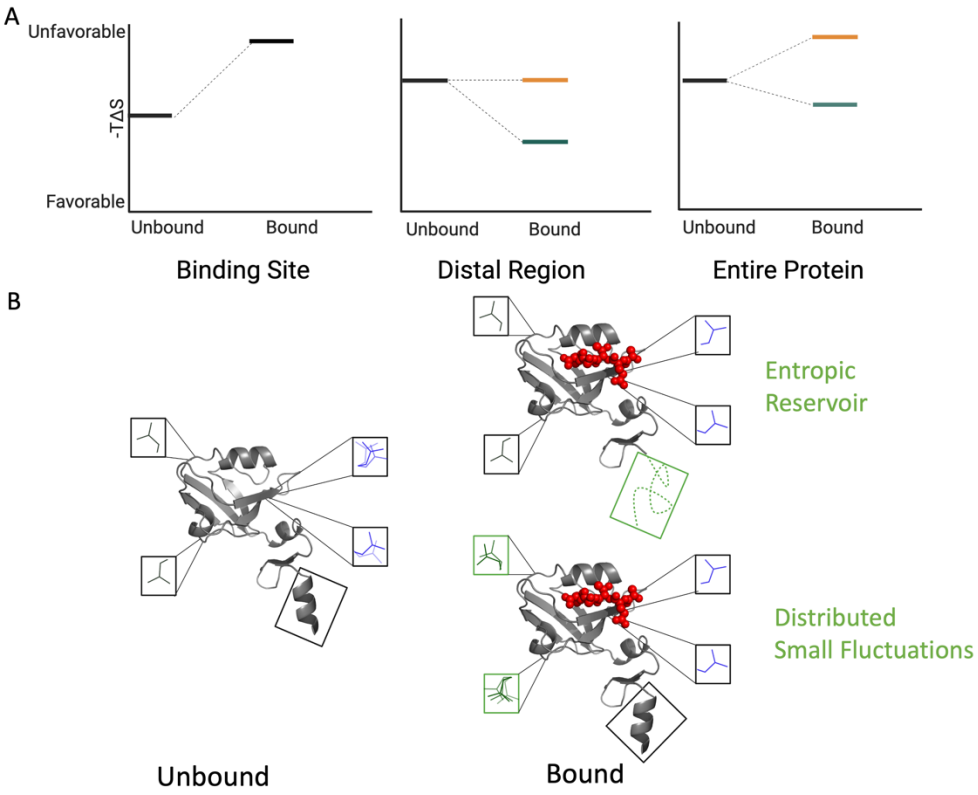

Figure 3. *(A) Upon binding, the binding site tends to lose conformational entropy due to hydrogen bonds and van der Waal interactions between the protein and ligand however, entropy can be redistributed to distal regions, leading to a increased conformational entropy, which can benefit binding affinity (teal). In contrast, if distal regions do not increase their entropy, the protein will play a larger entropic penalty that disfavors binding (orange).*

(B) Entropic redistribution upon binding can occur in many different ways. Upon binding the peptide (red), entropy can be redistributed from binding site residues (blue) to distal regions. This can occur through entropic reservoirs such as a region transition from ordered (alpha helix) to disordered (IDR). It can also occur through the globular part of the protein by increasing the entropy of residues distally, observed by residues accessing new rotameric states or increased harmonic disorder.

The most intuitive way of thinking about spatial redistribution of entropy is through portions of a protein acting as ‘entropic reservoirs’ that are poised to greatly change their conformational ensemble upon perturbation. Intrinsically disordered regions (IDRs) are prime candidates to serve as these entropic reservoirs. In human UDP-α-D-glucose-6-dehydrogenase (UGDH), the deletion of the C-terminal tail IDR reduces the affinity for the allosteric metabolite UDP-xylose by ~10-fold, despite making no direct contact with the ligand[44]. The ligand-bound state stabilizes altered contacts near the IDR, which allows the IDR to adopt an even more expanded conformational ensemble relative to the

apo state. Thus, a major driving force for binding is increased disorder and entropy in an already high-entropy IDR. To test this model, they showed that binding affinity lacked sequence dependence in the IDR region, but that the length of the IDR was strongly predictive of affinity. These results are consistent with an entropic force mode where the IDR can impact the globular protein region, and this impact is determined not only by the length of the IDR, but also by the underlying conformational ensemble of the IDR[45]. A similar interplay likely modulates allostery in the hTS example discussed above. In hTS, the interaction between an N-terminal IDR and the globular domains may influence the side chain dynamics underlying positive cooperativity. Intriguingly, this IDR segment is absent in the *E coli* homolog of TS, which is also dimeric but is non-cooperative[46]. In addition to IDRs, local unfolding provides another potential "entropic reservoir". In adenylate kinase, mutations that affect the local unfolding of lid subdomains are important for different aspects of catalysis, with an underlying entropic mechanism driving the changes in free energy[47]. These examples suggest that the modulation of thermodynamics by the entropy of unfolded or intrinsically disordered segments influences the dynamics of folded regions and can, therefore, tune the free energy of binding and allostery.

In addition to the large changes of the entropic reservoirs, small fluctuations within folded domains can contribute significantly to spatial entropic redistribution. The interplay between backbone and side chain in entropic spatial redistribution plays out in the relationship of peptide binding in PDZ domains. The third PDZ domain from PSD-95/SAP90 (PDZ3) shows a connection between a non-essential distal alpha helix and binding affinity. Truncation of alpha helix 3 (α3) reduces peptide binding affinity by 21-fold, even though it is over 6 Å away from the binding site. The lack of change in chemical shifts show this change is not due to a major structural perturbation change. Rather, the change is in the dynamics of the unbound state. In the wildtype protein, the side chains are much more rigid relative to the truncated variant. In the peptide bound state, both wildtype and the truncated variant have similar side chain dynamics. The change in binding affinity therefore results in the increased entropic penalty in the receptor. In this context, the α3 truncation is artificial, but it is a site where other domains are commonly linked to PDZ domains and may highlight a mechanism of inter-domain allosteric communication[48]. This type of interaction is reminiscent of the "entropic reservoirs" discussed above; however, in this case the dominant entropic contribution is on the side chain dynamics of the folded domain, not on the backbone of the terminal segment. The dominance of side chain dynamics in dictating the entropic contributions also played out in a study in the homologous second PDZ domain of human tyrosine phosphatase 1E. This PDZ domain demonstrates clear spatial redistribution: side chains around the binding site lose entropy while specific distal residues increase entropy[49]. As in the PDZ3 example, there is no evidence from chemical shifts that suggests a major conformational change. This result suggests that sub-angstrom changes in positions can be amplified non-linearly to enable alternative conformations of side chains at neighboring residues, greatly contributing to the number of states the protein populates. This is a provocative mechanism for a protein to have evolved ways to 'transfer' areas of higher entropy (i.e., more dynamics) through creating new voids and changes in rotamer states[8].

The most extensive evidence for this mechanism emerges from correlations between changes in NMR order parameters and calorimetry measurements. This correlation has been most extensively studied by Wand and colleagues in calmodulin, which binds various peptides with roughly similar overall binding affinity, but each with distinct entropic and enthalpic profiles. For example, upon

binding a peptide derived from smooth muscle myosin light chain kinase peptide (smMLCKp), the receptor loses side chain conformational entropy, especially around the smMLCKp binding site[50]. This result agrees with the direction and magnitude of the estimate of residual conformational entropy from calorimetry. In a series of papers, the authors extended this work to six peptides, each with similar affinity to calmodulin[20]. The contribution of conformational entropy to the binding affinity of these peptides varies widely, ranging from highly favorable to highly unfavorable. In each case, the average change in side chain NMR order parameters correlated with the estimates from calorimetry, leading to the concept of the 'entropy meter'[42] (**Figure 4**).

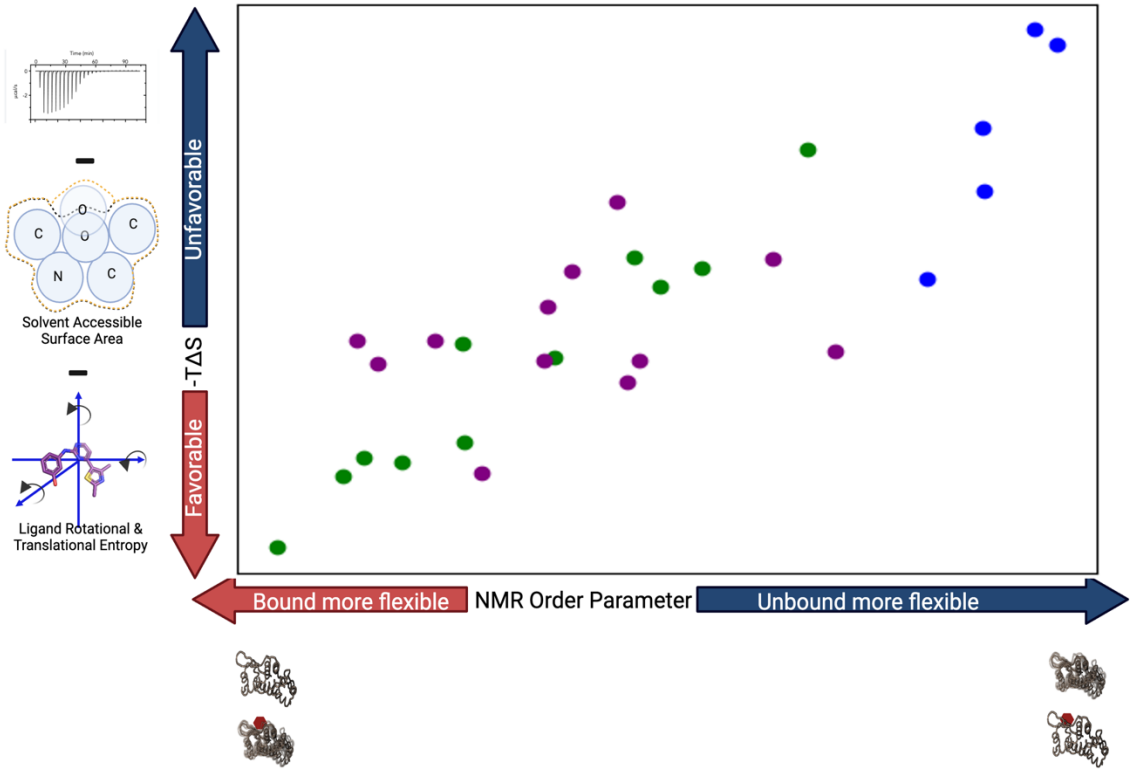

Figure 4. *Graphical representation of the entropy meter. On the x-axis is the delta order parameter of protein and ligands between the bound (top) and unbound (bottom) states. The y-axis represents the dS of the protein and ligand entropy measured by calorimeterty (estimates of solvent entropy was subtracted). Each point represents a unique protein-ligand complex and each color represents a different protein. Caro et al 2017 established that there is a linear relationship between the average delta NMR order parameter and calometric binding entropy, over a diverse set of protein and ligands, highlights the variability in conformational entropy changes upon binding. Proteins included: calmodulin (green data points), CAP (blue data points), Galectin, HEWL, PDZ3, PDZ2, DHFR, SAP (purple datapoints).*

This entropy meter represents the transformation of the dynamical indicator of conformational entropy into a quantitative metric. The entropy meter concept has now generalized to 28 protein-ligand complexes with eight different protein receptors[42]. Many of these examples do not have a clear spatial relationship of changes in entropy upon binding. However, it is difficult to discern if this is due to the sporadic methyl side chain probes used to determine the protein conformational entropy or the lack of spatial entropic redistribution in these systems. A spatial pattern is observed in the Catabolite activator protein (CAP), is a dimeric protein that binds DNA in a cAMP-dependent manner[51,52]. After cAMP binding, CAP undergoes a large conformational change to a state compatible with DNA-binding. A series of mutants with similar overall binding affinities have entropic contributions swing by almost 30 kcal/mol. Across these mutants, the ones with entropically driven binding have increased

side chain fluctuations in side chains distal from the DNA-binding site, while the residues near to DNA experience a loss of flexibility. The concept of how entropic contributions cluster in space has also been examined as a function of pressure in ubiquitin and for protein-protein interactions in barnase-barstar[53,54]. These studies also demonstrate clustering of side chains that increase or decrease flexibility in a correlated manner, suggesting that the redistribution of side chain dynamics may be an evolved feature of proteins to modulate the free energy of folding and binding.

To assess the general trends of entropic spatial redistribution after ligand binding, we identified ~700 pairs of high resolution X-ray crystal structures where the only difference was the presence of a small molecule ligand. We then re-built multiconformer models containing well-supported alternative conformations into the experimental data. These models allow us to estimate crystallographic order parameters[24] for each model and compare the conformational entropy change upon ligand binding[10]. A striking finding emerges from this analysis: the conformational entropy lost at the binding sites is correlated to the entropy gained at distal sites. We also observed diversity in the amount of entropy redistributed based on both the protein and ligand involved. Even for a single receptor, different ligands result in distinct patterns of entropy redistribution. Two ligands that cause equivalent reductions in entropy in the active site can differ by ~±1-2 kcals in the response at distal sites. This difference can occur because the ligands will engage different parts of the binding site in specific interactions that induce distinct patterns of spatial entropic redistribution. For example, in the kinase CDK2, the patterns of entropy in Type I inhibitors, which bind to the active-like conformation, are more similar to other Type I inhibitors than they are to Type II inhibitors, which bind to an inactive-like conformation. Similarly, Type II inhibitors were more similar to each other than they were to Type I inhibitors. However, even within the Type I inhibitors, ligands still showed unique patterns of conformational entropy, highlighting the possibility of designing for these features. The potential for tuning ligand properties to redistribute entropy also is evident across the entire dataset. Ligands with more specific and oriented interactions (e.g. hydrogen bonds) lead to greater local reductions in entropy and those binding events driven by apolar ligands lead to lower local reductions in entropy. Critically, these observations of spatial entropic redistribution are due to small, distributed changes in side chain alternative conformations and harmonic motions (B-factors), which are less well captured by traditional, single state structural models. This example highlights the critical importance of modeling the conformational ensemble to move towards a predictive understanding of the contributions of conformational entropy to ligand binding.

The importance of side chain entropy and the relevance of spatial redistribution to ligand binding extends beyond soluble proteins to membrane proteins. Many membrane proteins have high side chain conformational entropy in their apo state, which may offset the entropic penalty of sitting in the membrane[55]. The modulation of these dynamics upon ligand binding can influence binding and allostery[56]. For example, some GPCRs exhibit similar binding affinities for agonists and antagonists[57]. Multiple studies have identified how subtle changes in side chain conformational dynamics in the presence of either agonists or antagonists can influence the relative stability of specific conformations compatible with distinct binding downstream partners[58,59]. These results suggest a role for conformational entropy, located in specific distal regions of the protein, in controlling the free energy of binding and the output of allosteric signaling.

The concept of spatial entropic redistribution extends our thinking of the contributors to the free energy of binding to the rest of the scaffold. This strategic transfer of entropy is likely a fundamental aspect of dynamic allostery, where proteins distribute their inherent disorder in a controlled manner to facilitate allosteric regulation. There are multiple mechanisms that nature can harness to optimize the entropic contribution from distal regions. These include 'entropic reservoirs', such as intrinsically disordered regions (IDRs) and local unfolding, and subtle conformational changes to well packed regions poised to increase side chain dynamics. This mechanistic diversity emphasizes how the optimization of free energy can come from unexpected places and calls attention to the opportunity for considering entropy to predictably tune binding affinity.

Catalytically-Competent Ensembles

Enzymes catalyze chemical reactions through a high-energy transition state, where active site residues are precisely arranged to orient specific chemical groups responsible for catalysis. It follows that the positioning requirements will lead to an extreme reduction of conformational entropy for active site residues. How can this penalty be offset? One solution is by having these residues pre-organized in the apo or substrate-bound states. Another is the Circe effect, where the substrate binds in such a constrained, almost unique conformation, that is entropically disfavored[60]. A third potential strategy is for the enzyme to allow a much broader set of transition states, reducing the entropic penalty required in the binding site[61]. Finally, it is also possible to extend the concept of spatial redistribution for equilibrium ligand binding to catalysis.The capacity of residues in the rest of the enzyme ("the scaffold") to adopt many states that are compatible with the positioning required of catalytic residues throughout the catalytic cycle[62,63], can contribute favorably to reducing the entropic penalty of active site positioning. In the extreme case, if the scaffold conformational entropy is higher in the transition state than in the ground state, it can reduce the height of the barrier and entropically favor catalysis. This concept extends beyond the chemical step, which in most cases has a narrow transition state ensemble, to other steps of the catalytic cycle, each of which can be rate-limiting (e.g. substrate binding, product release, conformational change, etc). We call the ensemble of states compatible with the rate-limiting step of catalysis the *catalytically-competent ensemble* (**Figure 5A**).

Strong evidence for the importance of scaffold conformational entropy in catalysis emerges from a series of time-resolved crystallographic, NMR, and computational studies of the homodimeric enzyme, fluoroacetate dehalogenase (FaDH)[64,65]. Although the enzyme is dimeric, upon substrate binding at one active site, the other protomer does not bind substrate and can be thought of as an extended part of the scaffold. With clever mutations and substrate-analogues, these studies show how both solvent and protein scaffold entropy support the active site organization responsible for the chemical step of catalysis. This change to the scaffold occurs without any major backbone changes in any part of the catalytic cycle, but is evident in changes in: B-factors, alternative conformations of side chains, the number of bound waters, and NMR measurements of dynamics. Importantly, it is the contributions of distal residues and distal waters, not just waters displaced by the small molecules bound in the active site, that drive this entropic contribution. In addition, an analysis of MD trajectories using rigidity theory identified interprotomer allosteric pathways that suggests the mechanism of how these changes occur. Collectively these studies reveal how entropy can favor conformations responsible for rate-limiting steps of the catalytic cycle via increased conformational multiplicity of the scaffold and the release of solvent. The interplay between solvent and scaffold calls attention to the

importance of holistically modeling the conformational dynamics of the entire system, not just the active site.

The catalytically competent ensemble model can also be expanded to understand the operating principles of large macromolecular machines. For example, the ribosome undergoes many large structural rearrangements during the catalytic steps involved from initiation to elongation. These re-arrangements involve the making and breaking of many interactions that are relatively straightforward to interpret enthalpically. An elegant series of single-molecule fluorescence resonance energy transfer (smFRET) experiments leveraged temperature to probe the relative importance of enthalpy and entropy when two different charged tRNAs are bound to the ribosome. They revealed that both the equilibrium and rate of interconversion between the two conformational states of the ribosome, GS1 (associated with initiation) and GS2 (associated with elongation), have distinct entropic contributions depending on the identity of the tRNA present. The model that emerges is that the $tRNA^{Met}$, which is critically important for initiation, and $tRNA^{Phe}$, which only plays a role in elongation, differ in how they affect ribosome conformational dynamics. The energetic basis of this difference is driven by entropy favoring the corresponding states and barriers, indicating that the ribosome has evolved to manipulate entropy in response to specific tRNA-binding events that bias towards functionally relevant states (**Figure 5B**). How exactly this entropic difference is partitioned between broadening the transition state itself or an increase in the conformational entropy of the scaffold (including bound metals and water) remains an open question. This, and other studies demonstrating how the ribosome harnesses entropy during peptide bond formation[66] suggests a complex interplay of different entropic forces during its conformational and catalytic cycles of translation[67].

The catalytically-competent ensemble model also provides an energetic grounding for how conformational dynamics can be correlated with the evolvability of proteins. The relationship between increased conformational dynamics and evolvability is often discussed as altering the probability active sites conformational states[6], but it can extend to the energetic stabilization contributed by distal entropy in a catalytically competent ensemble. Promiscuous reactions, or errors, in catalysis are the first step towards evolutionary novelty and a broader conformational ensemble both locally in the active site and favored by distal conformational entropy can promote such reactions. Indeed, there is evidence for increased error rate for $tRNA^{Met}$ on the ribosome[68], which suggests that the entropically favoured transition state may also be broader at the active site. This model also provides a hypothesis for the contribution of some distal mutations accumulated in directed evolution experiments, especially for enzymes that change flexibility in regions beyond the active site[69,70]. With this framework, the challenge is to quantify how conformational heterogeneity in the protein scaffold (or, in the case of the ribosome, rRNA scaffold), substrate, and solvent molecules alter the conformational ensemble in a way that affects both the equilibrium and the transition rates between different functional steps of a catalytic cycle.

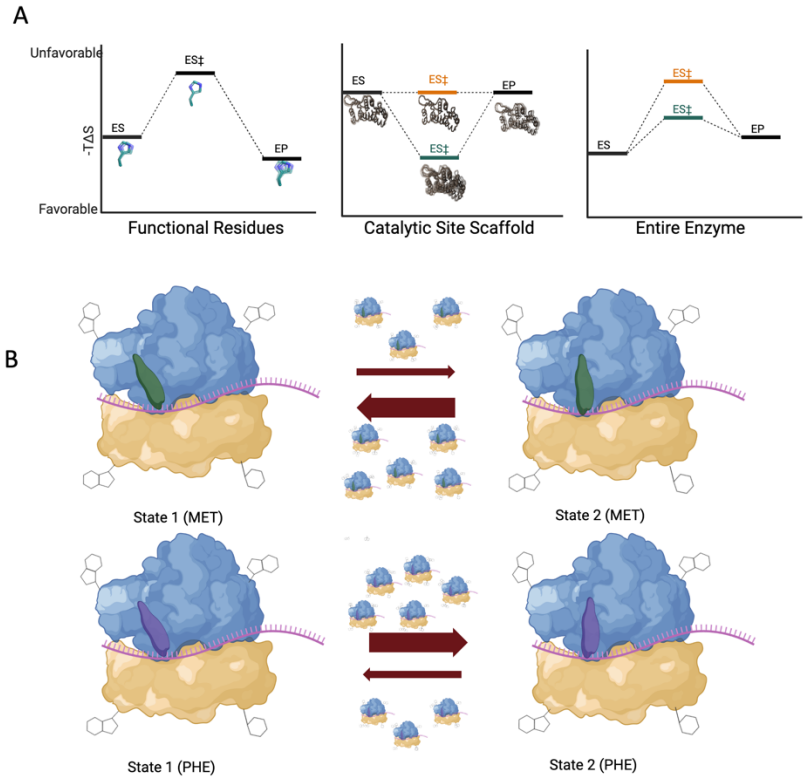

Figure 5. *(**A**) As the enzyme moves through the transition state, catalytic residues must undergo extreme entropic loss to allow for efficient chemistry. However, the catalytic site scaffold can still access many states, potentially increasing its entropy, even in the transition state, reducing the entropic loss across the entire enzyme (teal). This contrasts with the scenario where the enzyme scaffold also loses entropy, increasing the entropic loss across the entire enzyme (orange). (**B**) The ribosome (large subunit-blue; small subunit-yellow, mRNA-pruple) goes through large conformation changes between the initiation state (GS1) and elongation state (GS2). The more conformations the ribosome can access (represented by the number of ribosome states and size the red arrows), the more conformational entropy exists during the transition from GS1 to GS2 or vice versa.*

Future Directions in Discovering, Representing, and Manipulating Conformational Entropy

Failing to explicitly account for conformational entropy in structural biology skews our understanding of thermodynamics, hindering our ability to harness this source of free energy. Conformational entropy modulates free energy through three main mechanisms: *entropic pre-payment, spatial entropic redistribution,* and *the catalytically competent ensemble*. However, the lack of widespread structural ensemble representations limits our ability to explicitly explain the contributions of entropy and leverage this driving force for the prospective design of function[71]. Conformational ensembles, grounded in experimental data, will allow entropy to be integrated with probabilistic accounting of enthalpic-dominated interactions such as hydrogen bonds, salt bridges, and hydrophobic contacts to ground the complex interplay between structure, dynamics, and function in thermodynamics.

The most prominent link between conformational entropy and structural ensembles is inferred from NMR order parameters. However, there is ambiguity in the structural elements that give rise to a set of order parameters, as outlined in **Figure 1**. One way to resolve this ambiguity is to use experimental data as a bias in MD simulations[28,29]. This approach has provided insights into how mutations, counterintuitively, stabilize Chymotrypsin Inhibitor 2 through increased side chain dynamics in the

hydrophobic core[29]. The use of experimental biases in simulation is currently distinct from improved modeling and refinement methods that create structural ensemble models in X-ray crystallography or cryoEM [10,23,72–77]. Once refined, these models can simultaneously extract conformational entropy and structural interpretations[10], unhindered by the timescale constraints in NMR and MD simulations. Integrating multiple datasets simultaneously, particularly time-resolved[78] or fragment-soaking experiments[79], has the potential to further augment conformational entropy-structure discovery. The interface of experimental data, AlphaFold-type AI approaches, and physics-based molecular dynamics simulations will surely catalyze developments in quantifying conformational entropy[80,81]. Machine learning is already realizing some of this immense potential for ensemble discovery and model generation across length scales in CryoEM[82,83]. Understanding the relationship between latent spaces leveraged by these machine learning techniques and the underlying thermodynamics of the system may help guide the generative ensemble process to be even more predictive biologically. These methodological advancements are also necessitating better data structures than static PDB files to represent ensemble-level features[71]. Advancing the modeling and encoding of data will facilitate breakthroughs in structural ensemble predictions, enabling more facile interpretation of the conformational ensemble and how it relates to conformational entropy.

As we improve representations of other macromolecules it will be increasingly important to consider other sources of entropy, including solvent and ligands. Much attention has been focused on the behavior at the binding site, with reduction of the ligand ensemble into a single state being countered by the displacement of ordered water molecules. However, in high-resolution X-ray crystallography data, multiple conformations of ligands are often detected in binding sites, suggesting that many protein-ligand complexes have multiple binding states[84–86] and indicating that ligand binding results in less entropic loss than previously believed. Additionally, as seen in the FaDH model, solvent plays a critical role beyond the binding site. To gain a deeper insight into the entropic implications of solvent interactions, it is crucial to refine our modeling and thermodynamic interpretation of first and second shell water molecules, particularly those in exchange with the bulk solvent and thus 'partially occupied', as revealed by molecular dynamics (MD) simulations[87]. Creating a more atomistic model of the entropic contributions of solvent may reconcile outlier observations in the entropy meter, which links side chain dynamics to experimental ITC data, and move us closer to predicting the entropic impact of a perturbation.

The integration of structural and entropic features from conformational ensemble models will open up the use of a powerful lens to drive down ligand binding affinity and stabilize protein conformational states in design. Traditional drug design emphasizes optimizing the complementarity of a small molecule into the binding site. While this can improve enthalpy, it almost completely ignores entropy (except for consideration of the entropy gained by displaced solvent at the binding site). However, ligand properties, such as the number of hydrogen bonds, are associated with a difference in conformational entropy[10]. Further, minor chemical changes can significantly affect entropies and affinities, as seen in the hTS case where a single methyl group difference leads to vastly different entropies[19]. How minor changes to the ligand propagate to differences in protein conformational entropy remains a mystery (**Figure 6**). Structural perturbations do not necessarily propagate through structure via a series of dramatic “domino” conformational changes from one binding site to distal site. Instead, distributed and non-linear changes throughout the ensemble lead to changes in distal dynamics (entropy) or changes in the dominant conformational state (enthalpy). Examining these through a free energy landscape perspective, any conformational changes modifies its free energy

landscape and thus function, highlighting the need to combine, rather than isolate these features. Likewise, protein design, having been optimized for protein packing and stability, now faces the next frontier: designing improved binders and enzymes[88,89]. While pre-paying entropic cost by establishing order in the ground state has been explored[90], spatial entropic redistribution, achieved by increasing conformational entropy in functionally distal regions, is harder to design[91]. Integrating structural and entropic changes from natural versus designed enzymes could lead to innovative ways to design functional proteins.

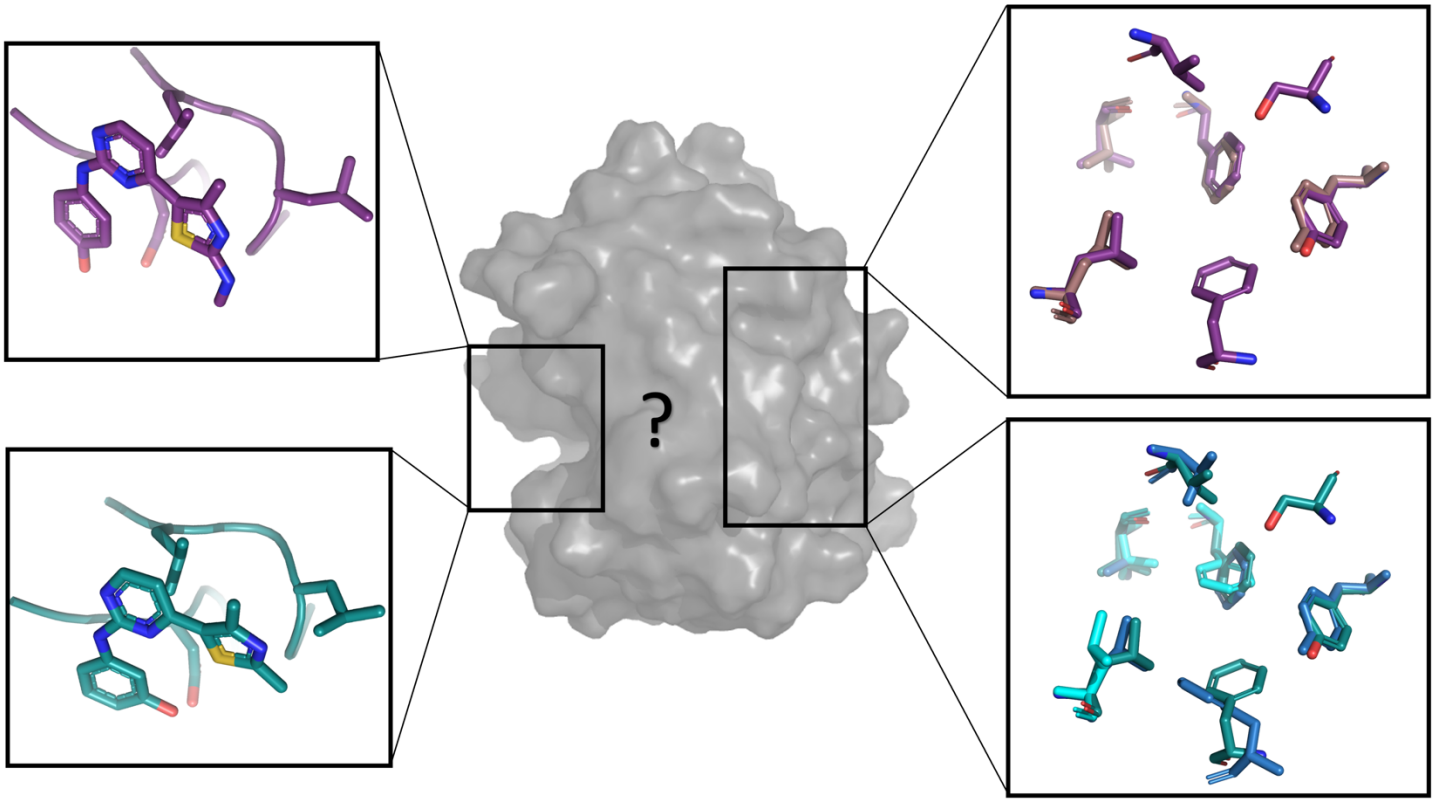

Figure 6. *Can spatial entropic redistribution can be leveraged in ligand design? Upon binding of two highly related ligands (differing by the position of a hydroxyl group and additional amine group), a binding site leucine changes rotamer states. Simultaneously, the conformational entropy of distal regions significantly differs between the two ligands. This leads the teal ligand to have significantly more conformational entropy in the distal region (represented by increased number of alternative conformations of distal residues), which could increase the teal ligand's binding affinity due to lower -TdS compared to the purple ligand. To use this prospectively, we need to understand the mechanism leading to increasing conformational entropy in distal regions.*

Collectively, the examples discussed here reveal distinct strategies discovered by nature to optimize the free energy of functionally important states. Although it is easier to explicitly represent and discuss specific and enthalpically directed interactions like hydrogen bonds, improved ensemble representations of macromolecular structure should allow for us to integrate structural and entropic information for perspective design of function. After all, nature cares about free energy, regardless of whether the source is easy to describe. Even if the energetic impact of conformational entropy is relatively small in some cases, the move from ensemble-averaged singular representations of protein structure to ensemble views could parallel the revolution catalyzed by single cell methods in genomics. In single cell work, the diversity of cell types and behaviors masked by bulk methods is leading to new insights into the properties of systems[92]. Indeed, the configurational entropy of cell

types may also play a role in tissue organization, connecting the importance of considering statistical mechanical concepts in decoding function across scales[93]. In the case of macromolecules, we look forward to future work that highlights how rare states contribute to specific functions and how conformational entropy can be leveraged as part of the toolkit to solve challenges in design, catalysis, and cellular signaling.

ACKNOWLEDGEMENTS

We thank Pei Zhou, Siyuan Du, David Sivak, Yifan Cheng, Margaux Pinney, Ruben Gonzalez, Shahar Sukenik, and members of the Fraser lab for their helpful comments. 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. Dill, K. A. Theory for the folding and stability of globular proteins. *Biochemistry* **24**, 1501–1509 (1985).
4. Doig, A. J. & Sternberg, M. J. Side-chain conformational entropy in protein folding. *Protein Sci.* **4**, 2247–2251 (1995).
5. 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).
6. Tokuriki, N. & Tawfik, D. S. Protein dynamism and evolvability. *Science* **324**, 203–207 (2009).
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. Bowman, G. R. & Geissler, P. L. Extensive conformational heterogeneity within protein cores. *J. Phys. Chem. B* **118**, 6417–6423 (2014).

9. Motlagh, H. N., Wrabl, J. O., Li, J. & Hilser, V. J. The ensemble nature of allostery. *Nature* **508**, 331–339 (2014).
10. 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).
11. 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).
12. Monod, J., Wyman, J. & Changeux, J. P. ON THE NATURE OF ALLOSTERIC TRANSITIONS: A PLAUSIBLE MODEL. *J. Mol. Biol.* **12**, 88–118 (1965).
13. Cooper, A. & Dryden, D. T. Allostery without conformational change. A plausible model. *Eur. Biophys. J.* **11**, 103–109 (1984).
14. 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).
15. 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).
16. Wand, A. J. & Sharp, K. A. Measuring Entropy in Molecular Recognition by Proteins. *Annu. Rev. Biophys.* **47**, 41–61 (2018).
17. 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).
18. 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).
19. Bonin, J. P., Sapienza, P. J. & Lee, A. L. Dynamic allostery in substrate binding by human thymidylate synthase. *Elife* **11**, (2022).
20. Marlow, M. S., Dogan, J., Frederick, K. K., Valentine, K. G. & Wand, A. J. The role of conformational entropy in molecular recognition by calmodulin. *Nat. Chem. Biol.* **6**, 352–358

(2010).

21. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).
22. 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).
23. 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.
24. 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).
25. 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).
26. 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).
27. 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).
28. Bottaro, S. & Lindorff-Larsen, K. Biophysical experiments and biomolecular simulations: A perfect match? *Science* **361**, 355–360 (2018).
29. Gavrilov, Y. *et al.* Double Mutant of Chymotrypsin Inhibitor 2 Stabilized through Increased Conformational Entropy. *Biochemistry* **61**, 160–170 (2022).
30. Yang, D. & Kay, L. E. Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding. *J. Mol. Biol.* **263**, 369–382 (1996).
31. Cui, Q. & Karplus, M. Allostery and cooperativity revisited. *Protein Sci.* **17**, 1295–1307 (2008).
32. Bonin, J. P. *et al.* Positive Cooperativity in Substrate Binding by Human Thymidylate Synthase.

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

33. Smith, C. A. *et al.* Population shuffling of protein conformations. *Angew. Chem. Int. Ed Engl.* **54**, 207–210 (2015).
34. 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).
35. 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).
36. Xie, T., Saleh, T., Rossi, P. & Kalodimos, C. G. Conformational states dynamically populated by a kinase determine its function. *Science* **370**, (2020).
37. Basanta, B. *et al.* The conformational landscape of human transthyretin revealed by cryo-EM. *bioRxiv* 2024.01.23.576879 (2024) doi:10.1101/2024.01.23.576879.
38. 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).
39. Kim, C. *et al.* A biophysical framework for double-drugging kinases. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2304611120 (2023).
40. Beyett, T. S. *et al.* Molecular basis for cooperative binding and synergy of ATP-site and allosteric EGFR inhibitors. *Nat. Commun.* **13**, 2530 (2022).
41. Borgia, A. *et al.* Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61–66 (2018).
42. Caro, J. A. *et al.* Entropy in molecular recognition by proteins. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 6563–6568 (2017).
43. Chodera, J. D. & Mobley, D. L. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu. Rev. Biophys.* **42**, 121–142 (2013).
44. Keul, N. D. *et al.* The entropic force generated by intrinsically disordered segments tunes protein function. *Nature* **563**, 584–588 (2018).

45. Yu, F. & Sukenik, S. Structural preferences shape the entropic force of disordered protein ensembles. *bioRxiv* (2023) doi:10.1101/2023.01.20.524980.
46. 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).
47. 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).
48. Lee, H.-J. & Zheng, J. J. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun. Signal.* **8**, 8 (2010).
49. Fuentes, E. J., Der, C. J. & Lee, A. L. Ligand-dependent dynamics and intramolecular signaling in a PDZ domain. *J. Mol. Biol.* **335**, 1105–1115 (2004).
50. 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).
51. Popovych, N., Sun, S., Ebright, R. H. & Kalodimos, C. G. Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* **13**, 831–838 (2006).
52. Tzeng, S.-R. & Kalodimos, C. G. Protein activity regulation by conformational entropy. *Nature* **488**, 236–240 (2012).
53. 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).
54. Fu, Y. *et al.* Coupled motion in proteins revealed by pressure perturbation. *J. Am. Chem. Soc.* **134**, 8543–8550 (2012).
55. 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).
56. 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).
57. Powers, A. S. *et al.* Structural basis of efficacy-driven ligand selectivity at GPCRs. *Nat. Chem.*

Biol. **19**, 805–814 (2023).

58. Clark, L. D. *et al.* Ligand modulation of sidechain dynamics in a wild-type human GPCR. *Elife* **6**, (2017).
59. Bumbak, F. *et al.* Ligands selectively tune the local and global motions of neurotensin receptor 1 (NTS). *Cell Rep.* **42**, 112015 (2023).
60. Jencks, W. P. Binding energy, specificity, and enzymic catalysis: the circe effect. *Adv. Enzymol. Relat. Areas Mol. Biol.* **43**, 219–410 (1975).
61. Jara, G. E. *et al.* Wide Transition-State Ensemble as Key Component for Enzyme Catalysis. *Elife* **12**, (2023).
62. 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).
63. Yabukarski, F. *et al.* Ensemble-function relationships to dissect mechanisms of enzyme catalysis. *Sci Adv* **8**, eabn7738 (2022).
64. Kim, T. H. *et al.* The role of dimer asymmetry and protomer dynamics in enzyme catalysis. *Science* **355**, (2017).
65. Mehrabi, P. *et al.* Time-resolved crystallography reveals allosteric communication aligned with molecular breathing. *Science* **365**, 1167–1170 (2019).
66. 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).
67. Kaledhonkar, S. *et al.* Late steps in bacterial translation initiation visualized using time-resolved cryo-EM. *Nature* **570**, 400–404 (2019).
68. Gamper, H. B., Masuda, I., Frenkel-Morgenstern, M. & Hou, Y.-M. Maintenance of protein synthesis reading frame by EF-P and m(1)G37-tRNA. *Nat. Commun.* **6**, 7226 (2015).
69. Chao, F.-A. *et al.* Structure and dynamics of a primordial catalytic fold generated by in vitro evolution. *Nat. Chem. Biol.* **9**, 81–83 (2013).

70. Campbell, E. *et al.* The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol.* **12**, 944–950 (2016).
71. Wankowicz, S. & Fraser, J. Comprehensive encoding of conformational and compositional protein structural ensembles through mmCIF data structure. *ChemRxiv* (2023) doi:10.26434/chemrxiv-2023-ggd1w-v2.
72. Ginn, H. M. Torsion angles to map and visualize the conformational space of a protein. *Protein Sci.* **32**, e4608 (2023).
73. 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).
74. 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).
75. Burnley, B. T., Afonine, P. V., Adams, P. D. & Gros, P. Modelling dynamics in protein crystal structures by ensemble refinement. *Elife* **1**, e00311 (2012).
76. Beton, J. G., Mulvaney, T., Cragnolini, T. & Topf, M. Cryo-EM structure and B-factor refinement with ensemble representation. *Nat. Commun.* **15**, 444 (2024).
77. Bonomi, M. & Camilloni, C. Integrative structural and dynamical biology with PLUMED-ISDB. *Bioinformatics* **33**, 3999–4000 (2017).
78. Pearson, A. R. & Mehrabi, P. Serial synchrotron crystallography for time-resolved structural biology. *Curr. Opin. Struct. Biol.* **65**, 168–174 (2020).
79. Pearce, N. M. *et al.* A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat. Commun.* **8**, 15123 (2017).
80. Hoff, S. E., Thomasen, F. E., Lindorff-Larsen, K. & Bonomi, M. Accurate model and ensemble refinement using cryo-electron microscopy maps and Bayesian inference. *bioRxiv* 2023.10.18.562710 (2023) doi:10.1101/2023.10.18.562710.
81. Vani, B. P., Aranganathan, A., Wang, D. & Tiwary, P. AlphaFold2-RAVE: From Sequence to Boltzmann Ranking. *J. Chem. Theory Comput.* **19**, 4351–4354 (2023).

82. Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat. Methods* **18**, 176–185 (2021).
83. 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).
84. Skaist Mehlman, T. *et al.* Room-temperature crystallography reveals altered binding of small-molecule fragments to PTP1B. *Elife* **12**, (2023).
85. 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).
86. 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).
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. 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).
89. 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).
90. St-Jacques, A. D. *et al.* Computational remodeling of an enzyme conformational landscape for altered substrate selectivity. *Nat. Commun.* **14**, 6058 (2023).
91. Dagliyan, O. *et al.* Engineering extrinsic disorder to control protein activity in living cells. *Science* **354**, 1441–1444 (2016).
92. Quake, S. R. A decade of molecular cell atlases. *Trends Genet.* **38**, 805–810 (2022).
93. Srivastava, V. *et al.* Configurational entropy is an intrinsic driver of tissue structural heterogeneity. *bioRxiv* (2023) doi:10.1101/2023.07.01.546933.