Received: 4 May 2022 | Revised: 5 August 2022 | Accepted: 9 August 2022

DOI: 10.1002/pro.4424

FULL-LENGTH PAPER

THE PROTEIN SOCIETY WILEY

Thermodynamic coupling between neighboring binding sites in homo-oligomeric ligand sensing proteins from mass resolved ligand-dependent population distributions

Weicheng Li[1] | **Andrew S. Norris**[1,2] | **Katie Lichtenthal**[3] | **Skyler Kelly**[3] | **Elihu C. Ihms**[4] | **Paul Gollnick**[3] | **Vicki H. Wysocki**[1,2] | **Mark P. Foster**[1]

[1]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio, USA

[2]Resource for Native Mass Spectrometry Guided Structural Biology, The Ohio State University, Columbus, Ohio, USA

[3]Department of Biological Sciences, University at Buffalo, State University of New York, Buffalo, New York, USA

[4]Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA

Correspondence
Vicki H. Wysocki and Mark P. Foster, Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210, USA.
Email: wysocki.11@osu.edu, foster.281@osu.edu

Funding information
National Institute of General Medical Sciences, Grant/Award Numbers: GM077234, GM120923, GM128577; National Institutes of Health, Grant/Award Numbers: P41 GM128577, R01 GM077234, R01 GM120923

Reviewing Editor: John Kuriyan

Abstract

Homo-oligomeric ligand-activated proteins are ubiquitous in biology. The functions of such molecules are commonly regulated by allosteric coupling between ligand-binding sites. Understanding the basis for this regulation requires both quantifying the free energy ΔG transduced between sites, and the structural basis by which it is transduced. We consider allostery in three variants of the model ring-shaped homo-oligomeric *trp* RNA-binding attenuation protein (TRAP). First, we developed a nearest-neighbor statistical thermodynamic binding model comprising microscopic free energies for ligand binding to isolated sites ΔG_0, and for coupling between adjacent sites, ΔG_α. Using the resulting partition function (PF) we explored the effects of these parameters on simulated population distributions for the 2^N possible liganded states. We then experimentally monitored ligand-dependent population shifts using conventional spectroscopic and calorimetric methods and using native mass spectrometry (MS). By resolving species with differing numbers of bound ligands by their mass, native MS revealed striking differences in their ligand-dependent population shifts. Fitting the populations to a binding polynomial derived from the PF yielded coupling free energy terms corresponding to orders of magnitude differences in cooperativity. Uniquely, this approach predicts which of the possible 2^N liganded states are populated at different ligand concentrations, providing necessary insights into regulation. The combination

Weicheng Li and Andrew S. Norris contributed equally to this study.

of statistical thermodynamic modeling with native MS may provide the thermodynamic foundation for a meaningful understanding of the structure–thermodynamic linkage that drives cooperativity.

KEYWORDS

cooperativity, homo-oligomer, native MS, statistical thermodynamics

1 | INTRODUCTION

1.1 | Homotropic allostery is widespread in biology

Homotropic cooperativity describes a ubiquitous process in biological regulation. It arises commonly in proteins that form symmetric homo-oligomeric assemblies whose biological activity is modified by ligands.[1,2] In such symmetric assemblies, even though the binding sites are identical, ligand-binding affinities depend on the liganded state of the oligomeric protein; that is, the affinity K of a ligand for a receptor with no bound ligands will be different for a receptor that already has one or more bound ligands.[3] This difference in affinity arises through "allosteric" interactions between the ligand-binding sites, which alter the free energy change upon additional ligand binding.[4,5] Such thermodynamic changes are most often associated with structural changes that accompany ligand binding, though changes in protein dynamics have also been implicated.[6–8] A large proportion of biological processes are regulated by oligomeric ligand-binding proteins, from hemoglobin[9] to p97 ATPase.[10,11] Therefore, to understand biological regulation we must first understand the linkage between ligand-induced structural changes and changes to protein-ligand free energy landscapes.[12]

1.2 | TRAP protein rings as models of homotropic allostery

We explore homotropic cooperativity in the homo-oligomeric *trp* RNA-binding attenuation protein (TRAP), which plays a central role in regulating tryptophan (Trp) production in Bacilli.[13] TRAP was originally identified in *Bacillus subtilis* (*Bsu*) as a 76-residue protein that assembles into homo-oligomeric complexes. Upon binding multiple Trp ligands, TRAP is activated to bind specific RNA sequences in the 5′ untranslated region of the *trp* operon mRNA.[14,15] Crystal structures of TRAP from *Bsu*, *B. stearothermophilus* (*Bst*), *B. licheniformis*, and *B. halodurans* (*Bha*) revealed undecameric (11mer) and dodecameric (12mer) rings with binding sites for Trp in the interfaces between each of the protomers (Figure 1a–c). Trp-dependent sequence-specific RNA recognition is achieved via base contacts from amino acids located on loops that are shown by NMR studies to be flexible in the absence of Trp and structured upon Trp binding[16,17]; these loops thus mediate heterotropic Trp-RNA cooperativity. The same loops that mediate RNA binding also are points of contact between neighboring Trp-binding sites, providing a structural scaffolding for mediating site-site interactions (Figure 1d).

1.3 | Quantifying the detailed thermodynamics of ligand binding to homo-oligomeric proteins is difficult

Experimentally, binding affinities are usually quantified by measuring ligand-dependent changes in an observable that is proportional to the fraction of bound states Y; for example, a spectroscopic signal, or heat released upon binding. For binding a ligand L to its site on a macromolecule M to form an ML complex, $M + L \rightleftharpoons ML$, Y has the familiar hyperbolic form $Y = \frac{[L]}{[L]+1/K}$. Here [L] is the free ligand concentration, and the equilibrium association constant K is related to the free energy difference between the free and bound states, $\Delta G = -RT\ln K$. The population of the bound state is proportional to its Boltzmann probability, $p_{\text{bound}} \propto \exp(-\Delta G/RT)$. So, measuring a signal proportional to Y at different ligand concentrations [L] yields K and the binding free energy ΔG.

For oligomeric macromolecules, the measured signal is averaged over multiple sites and configurations. For a dodecameric protein with 12 binding sites, there are $2^{12} = 4{,}096$ possible configurations with 0–12 bound ligands. If the binding sites are independent and have the same affinity K for the ligand (i.e., in the absence of cooperativity), the most likely configurations are those with the greatest statistical degeneracy: the populations of states with 0–12 bound ligands L follow a binomial distribution: $P(n, M, Y) = \frac{M!}{n!(M-n)!} Y^n (1-Y)^{M-n}$, where n is the number of bound ligands, M is the total number of sites; Y, the fraction of bound sites, is equivalent to the

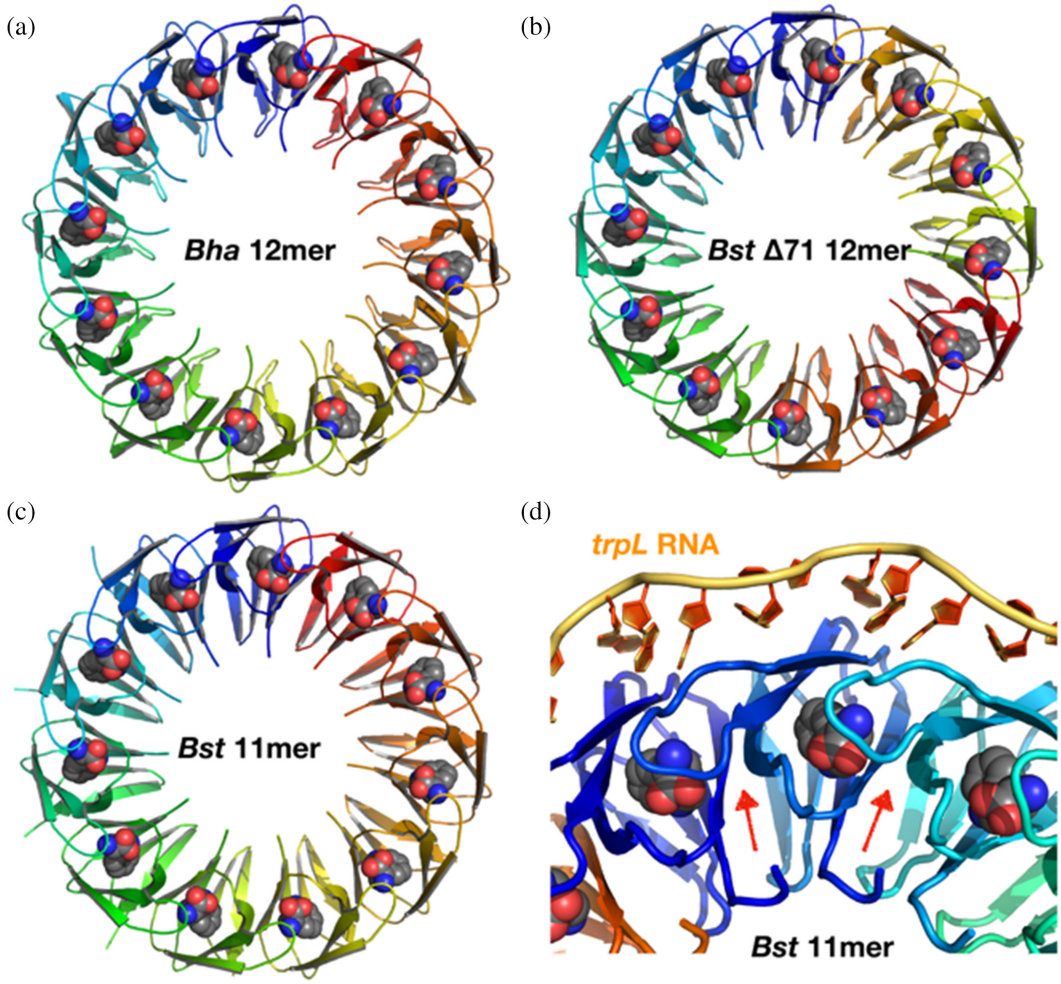

FIGURE 1 TRAP forms ring-shaped homo-oligomers to bind Trp and RNA. (a–c) Crystal structures of Trp-bound TRAP revealed 12mers for *Bha* TRAP (a; 3zzl), and for an engineered *Bst* TRAP Δ71 (b; 3zzs), while wild-type *Bst* TRAP crystallized as 11mers (c; 1qaw). Bound Trp ligands are shown as spheres. (d) Crystal structure of *Bst* TRAP bound to an idealized trp leader RNA (green; 1c9s). Protein loops (red arrows) that are disordered in Trp-free TRAP and become structured upon Trp binding, participate in RNA recognition and mediate allosteric interactions between Trp-binding sites. Each TRAP ring can adopt 2^N (=2,048 or 4,096) states with 0–*N* bound Trp where *N* equals 11 or 12. Measuring the populations of these states is required to quantify magnitude and mechanism of allostery. TRAP, *trp* RNA-binding attenuation protein; Trp, tryptophan

probability of sites being bound.[18,19] Thus, for example, when enough ligand is present so that on average one-twelfth of all sites are occupied ($Y = 1/12 = 0.083$), states with one ligand bound are only slightly more abundant than those with none bound, or with two bound (0: 35%, 1: 38%, 2: 19%, 3: 6%) (Figure S1). If enough ligand is present to fill half of the sites ($Y = 0.5$), the most dominant states are those with six bound ligands because there are simply more ways to arrange 6 ligands on a 12-site lattice, $\frac{M!}{n!(M-n)!} = \frac{12!}{6!6!} = 924$. Under these conditions the states with 0–12 bound ligands are populated according to a *normal* distribution, with coefficients of 1:12:66:220:495:792:924:792:495:220:66:12:1 (Figure 2). The experimentally measured signal is thus the average over all liganded states and their configurations. Nevertheless, if the sites are identical and independent, and the signal is proportional to *Y*, we can determine *K* by measuring how it changes with [L].

Cooperativity implies that ligand-binding events to otherwise identical sites occur with different affinities. Because affinities and free energy changes are linked, these differences result in altered population distributions. At the extreme of infinitely positive cooperativity, under conditions of half maximal saturation ($Y = 0.5$), instead of a *normal* distribution we would observe a strict bimodal distribution: half of the oligomers will be fully bound, and half will be empty. Thus, cooperativity makes some configurations more favorable than others; which states are more favored depends on the *mechanism* and *magnitude* of the cooperative interactions. Thus, the

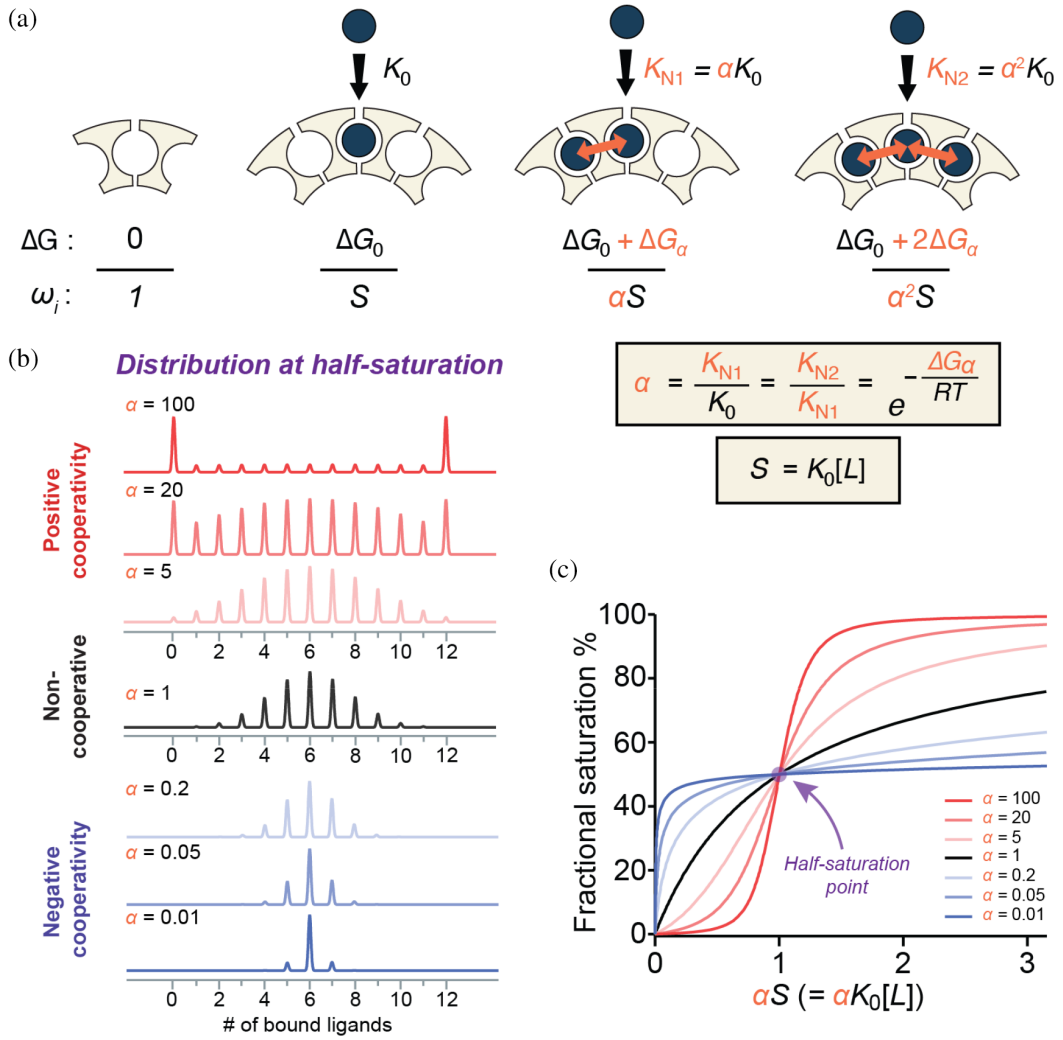

FIGURE 2 Nearest-neighbor models can explain and predict population distributions for cooperative ligand binding to ring-shaped oligomeric proteins. (a) In an NN model of cooperativity, four basic ligand-binding modes can be described with two energy parameters: an intrinsic free energy change for binding to sites with no occupied neighbors (ΔG_0), and a coupling free energy term contributed solely from each adjacent bound ligand (ΔG_α). An empty site is defined as the reference state with $\Delta G_0 = 0$ and a statistical weight of 1. Binding to an isolated site is afforded a statistical weight $\omega = S = K_0[\mathrm{L}]$. Cooperativity from one bound neighbor results in an association constant $K_{N1} = \alpha K_0$, $\omega = \alpha S$, while binding to a site with two occupied neighbors is occurs with affinity $K_{N2} = \alpha^2 K_0$, and $\omega = \alpha^2 S$. α values >1 indicate stronger binding to sites with occupied neighbors. The four liganded states are referred to in the text as 0, N0, N1 and N2, respectively. (b) Effect of NN cooperativity on simulated population distributions for dodecameric rings with 0–12 bound ligands at half-saturation. Simulated populations for a non-cooperative binding ($\alpha = 1$, black), equivalent to non-interacting sites, yield the expected binomial distribution. Population distributions arising from both positive ($\alpha > 1$, red) and negative ($\alpha < 1$, blue) cooperativity were simulated using a cyclic NN model. In the presence of stronger positive cooperativity the apo and holo populations dominate the distribution. In contrast, negative cooperativity favors states with half of the sites occupied. (c) Effects of NN cooperativity on fractional saturation of ligand-binding sites. Positive cooperativity (red) generates sigmoidal binding curves, the non-cooperative binding curve is hyperbolic (black), whereas negative cooperativity (blue) generates curves that are steeper at low concentration but shallow at high concentration. NN, nearest-neighbor

population distribution encodes rich information about cooperative mechanisms that is largely lost when convolved into the average metric *Y*.

A complicating factor in measuring population distributions is that cooperativity may distort the proportionality between the observable signal and fractional saturation *Y*. For example, in isothermal titration calorimetry (ITC) experiments, ligands may bind with differing affinities K_i, but also with different enthalpies ΔH_i; this complicates the use of heat as a metric for quantifying populations. Likewise, if a ligand bound at a site allosterically favors structural shifts in neighboring binding

sites, monitoring subsequent ligand-binding events by circular dichroism (CD) might generate smaller structural changes and therefore lose proportionality to bound population. Thus, it can be difficult to determine populations of states from typical bulk-averaged observables.

1.4 | Native mass spectrometry can measure population distributions

Native mass spectrometry (MS) has the potential to distinguish allosteric mechanisms by more directly probing bound-state distributions.[7,8,20–22] Unlike measurements of average fractional saturation, native MS can directly probe the populations of states with different numbers of bound ligands.[22] The mass (and mass-to-charge ratios, m/z) of a macromolecular complex is exactly determined by the molecular weight of its constituents. Thus, for an ionized macromolecule with n bound ligands, the signal will appear at the m/z dictated by the mass of the components and the charge on the ion. This direct proportionality between signal and ligand occupancy thus makes native MS well suited to characterize populations of states for molecules with multiple binding sites.

1.5 | Proper model selection is important

Finally, a severe confounding factor in understanding mechanisms of cooperativity is the selection of an appropriate model to extract thermodynamic parameters from binding data. In the absence of cooperativity, one may fit the signal change with the hyperbolic binding equation $\Delta\text{signal} = \text{amplitude} \times \frac{[\text{L}]}{[\text{L}]+1/K}$ if the free ligand concentration is known or can be approximated, or the quadratic equation:

$$\Delta\text{signal} = \text{amplitude} \times \frac{[\text{L}]_\text{T} + [\text{M}]_\text{T} + 1/K - \sqrt{\left([\text{L}]_\text{T} + [\text{M}]_\text{T} + 1/K\right)^2 - 4[\text{L}]_\text{T}[\text{M}]_\text{T}}}{2[\text{M}]_\text{T}}, \quad (1)$$

if only total ligand and binding site concentrations $[\text{L}]_\text{T}$ and $[\text{M}]_\text{T}$ are known; amplitude is the proportionality between fractional saturation Y and the signal change. However, these equations are inappropriate for cooperative binding since they only include one equilibrium constant K. Cooperativity is plainly evident if binding curves cannot be adequately described by one of these equations.

Parsimonious approaches to data analysis demand adding only as many parameters as needed to produce a good fit to the binding data. A common approach is to use the Hill equation, which can generate sigmoidal binding curves but is uninformative about free energy differences between binding modes.[23] Phenomenological multi-site independent binding models can also produce cooperative binding curves as well as different apparent affinities, yielding free energy differences between phenomenological binding modes.[24]

However, to decipher cooperativity mechanisms, it is necessary to apply an appropriate physical binding model to the thermodynamic parameters. In the concerted Monod-Wyman-Changeux (MWC) model, all subunits of a homo-oligomer are posited to simultaneously exchange between distinct structures with different affinities for the ligand binding.[25] In the sequential Koshland-Nemethy-Filmer (KNF) model, each subunit independently undertakes conformational changes upon ligand binding, altering the binding affinity for subsequent binding events.[26] While each of these models are able to reproduce cooperative binding curves, they are unable to clarify which of the $2^N = 4{,}096$ possible liganded configurations of a 12mer are more favorable, or how ligand-binding events alter the free energies for other binding events.

1.6 | Mechanistic insights from statistical thermodynamic models and native MS

Here, we explore the use of mechanistic nearest-neighbor (NN) lattice-based statistical thermodynamic models, along with native MS, to quantify and describe the cooperative behavior of three TRAP protein variants. We apply an NN model parametrized with the free energy of binding to isolated sites ΔG_0, and a microscopic free energy for communication between sites ΔG_α. We iterate over all possible 2^N liganded configurations of TRAP rings to develop a statistical thermodynamic partition function (PF) that describes the ligand-dependent probabilities of each of the 2^N possible states. We explore the effect of NN interaction energies on population distributions during simulated ligand titrations and examine the sensitivity of average site occupancy to the underlying population distributions.

We then measure Trp binding to TRAP variants using conventional bulk methods (CD, ITC), and using native MS. We find that native MS makes it possible to resolve and quantify states with 0–N bound ligands, thereby allowing measurement of the statistical thermodynamic PF. Fitting of the populations of states with different numbers of bound ligands to a binding polynomial derived from the PF allows quantification of the

microscopic free energy transduced between adjacent ligand-binding sites. This is the thermodynamic quantity required for a true understanding of structure–thermodynamic relationships in biological regulation.

2 | RESULTS

2.1 | Nearest-neighbor statistical thermodynamic model can quantify cooperativity and provide mechanistic insights

A NN model posits that the thermodynamics of ligand binding to a site are influenced by the liganded states of the immediately adjacent sites.[5] Thus, the thermodynamics of binding to a given site can be parametrized using two microscopic free energy terms (ΔG_0, ΔG_α, Figure 2a), corresponding to free energies for binding to sites with no bound neighbors, and thermodynamic contributions of interactions with occupied neighbors, respectively; $\Delta G_\alpha = 0$ in the absence of cooperativity. Each of the N sites in the macromolecule can be classified as either: (a) empty, state-0 ($\Delta G = 0$, the reference state); (b) bound with no occupied neighbors, state-N0 (ΔG_0); (c) bound with one occupied neighbor, state-N1 (ΔG_0 +- ΔG_α); or (d) state-N2 ($\Delta G_0 + 2\Delta G_\alpha$), bound with two occupied neighbors. The corresponding microscopic ligand-binding affinity constants for each state are $K_0 = 1$, $K_{N0} = \exp\left(-\frac{\Delta G_0}{RT}\right)$, $K_{N1} = \exp\left(-\frac{\Delta G_0 + \Delta G_\alpha}{RT}\right)$, and $K_{N2} = \exp\left(-\frac{\Delta G_0 + 2\Delta G_\alpha}{RT}\right)$, respectively. In addition, it is convenient to quantify "cooperativity" by comparing the affinities for ligand binding in various modes, such that $\alpha = K_{N1}/K_0 > 1$ quantifies the fold *increase* in affinity afforded by having the first occupied neighbor, while $\alpha < 1$ quantifies the fold *decrease* in affinity from negative cooperativity. Likewise, α^2 is the fold change in affinity to a site with two bound neighbors, compared to none. An $\alpha = 10$ means a ligand binds 10 times more tightly to a site with one occupied neighbor than to one with none, and 100 times more tightly to a site with two flanking neighbors; it also corresponds to a thermodynamic coupling between sites of $\Delta G_\alpha = -RT\ln 10 = -2.3RT$, or $-1.4\,\text{kcal mol}^{-1}$ at 298 K. With these parameters, NN models enable elaboration of a statistical partition function that enumerates the free energies and Boltzmann probabilities of all possible 2^N states by straightforward iteration and examination of their NN site occupancy.[5,39]

Application of an NN model to a cyclic homo-dodecamer results in a partition function (PF) Ξ_{NN} with 38 unique energetic terms corresponding to the summed statistical weights ω_i for the $2^{12} = 4{,}096$ configurations:

$$\begin{aligned}\Xi &= 1\\ &+S(12)\\ &+S^2(54+12\alpha)\\ &+S^3(112+96\alpha+12\alpha^2)\\ &+S^4(105+252\alpha++126\alpha^2+12\alpha^3)\\ &+S^5(36+240\alpha+360\alpha^2+144\alpha^3+12\alpha^4)\\ &+S^6(2+60\alpha+300\alpha^2+400\alpha^3+150\alpha^4+12\alpha^5)\\ &+S^7(36\alpha^2+240\alpha^3+360\alpha^4+144\alpha^5+12\alpha^6)\\ &+S^8(105\alpha^4+252\alpha^5+126\alpha^6+12\alpha^7)\\ &+S^9(112\alpha^6+96\alpha^7+12\alpha^8)\\ &+S^{10}(54\alpha^8+12\alpha^9)\\ &+S^{11}\alpha^{10}(12)\\ &+S^{12}\alpha^{12}\end{aligned} \quad (2)$$

Here, $S = K_0[L]$, where [L] is the free Trp concentration and K_0 and α are as in Figure 2; generation of the PF was automated using *itcsimlib*.[5] Each row in Equation (2) corresponds to states with different numbers of bound ligands, from 0 to 12, with 1 assigned as the statistical weight of the state with zero bound ligands. For each term, the coefficients reflect the degeneracy of the energetic states. For example, there are 12 equivalent ways to configure 1 or 11 ligands on a 12-mer. For two bound ligands, there are 12 ways to arrange them next to each other, with statistical weights modified by the cooperativity factor α, and 54 ways to arrange them in isolation, with no interaction term (Figure S3A). The exponent on the α terms reflects the number of NN interactions in that configuration, such that, for example, the half-bound state with the $400\alpha^3$ term corresponds to configurations that arrange the six bound ligands with three pairs of interactions (e.g., Figure S3B).

With the NN partition function constructed it is possible to predict the effect of cooperativity on the populations of each of the 4,096 states in a ring-shaped homo-dodecameric proteins. As noted above, in the absence of cooperativity, at ligand concentrations that result in half of the sites being bound, the populations follow a binomial distribution (Figure 2b, black). Positive cooperativity increases the population of states with more bound ligands because they feature more NN interactions; this also results in more states with fewer sites bound, due to mass balance (Figure 2b, red; Figure S4A). In the case of strong positive cooperativity, apo and holo states dominate the population distribution, with intermediate states being lowly populated. In contrast, negative cooperativity has the opposite effect, favoring half-bound configurations that spread out ligands over all available sites while minimizing NN interactions (Figure 2b, blue; Figure S4B).

In addition to providing details of the underlying populations, NN models can be used to compute bulk properties for comparison with experimental measurements (Figure 2c). A plot of fractional saturation versus free ligand concentration generates the expected hyperbolic curve in the absence of cooperativity (black). Positive cooperativity produces the expected sigmoidal curves (red) and negative cooperativity yields compressed saturation curves (blue). For reference, we also fit the trajectories with the Hill equation. This yielded Hill coefficients of ~8 and ~0.2 for $\alpha = 100$, and 0.01, respectively. However, while α is directly related to the coupling-free energy, the Hill coefficient is phenomenological and not mechanistically informative.

2.2 | Cooperative Trp binding to multimers can be difficult to characterize using traditional bulk binding assays

To quantify Trp binding to oligomeric TRAP rings, we first used CD spectropolarimetry to measure its ellipticity at 228 nm, which reports on Trp binding (Figure 3a).[31] CD measurements were conducted at 25°C with ~24 μM

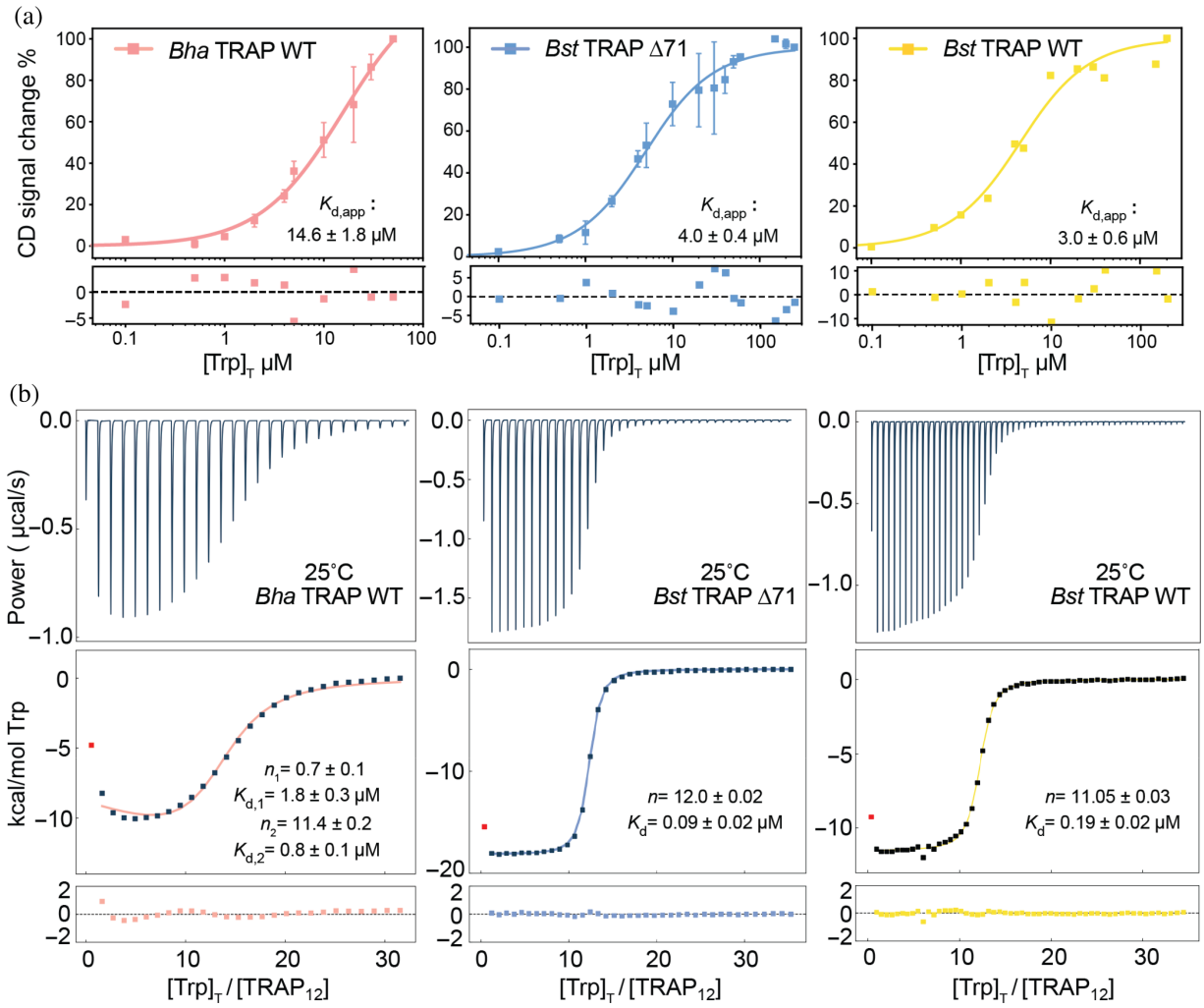

FIGURE 3 Bulk binding and phenomenological modeling provide limited information on cooperative ligand-binding to dodecameric *Bha* TRAP and *Bst* TRAP Δ71 (left, middle), or undecameric wild-type *Bst* TRAP (right). (a) The circular dichroism signal at 228 nm in 50 mM sodium phosphate buffer (pH = 8) was measured as a function of added Trp, and the data fit with the quadratic binding equation (Equation (1); Table S1). CD data were acquired by using 2 μM TRAP with increasing Trp concentration from 0 to 300 μM. The fitted parameters from CD suggest all three proteins exhibit similar μM-binding affinity and homotropic cooperativity is not evident. (b) ITC experiments titrating Trp into TRAP at 25°C in buffer A (see Methods) were fit with quadratic independent site models (Table S2). *Bst* TRAP Δ71 (middle) and *Bst* TRAP (right) can be well fitted with one-site binding model, in which 12mer *Bst* TRAP Δ71 ($K_d = 0.09 \pm 0.02$ μM) has two-times higher binding affinity than 11mer *Bst* TRAP ($K_d = 0.19 \pm 0.02$ μM). The data for *Bha* TRAP (left) are better fit to a two-sites binding model yielding parameters in which the first Trp binds with K_d of 1.8 ± 0.3 μM while the remaining 11 bind with a K_d of 0.8 ± 0.1 μM. CD, circular dichroism; TRAP, *trp* RNA-binding attenuation protein

TRAP (2 μM rings) over a range of Trp concentrations. Binding curves were fit with the binding quadratic (Equation (1)), yielding apparent dissociation constants ($K_{d,app} = 1/K_{app}$). The three TRAP variants yielded similar apparent Trp-binding affinities of 3–14 μM (Table S1), consistent with previous findings.[27] Under these conditions, cooperative binding of Trp to the TRAP variants is not evident.

Homotropic cooperativity can alter thermodynamic parameters other than binding-free energy (e.g., ΔH), so we also studied Trp binding to TRAP using isothermal titration calorimetry (ITC), which directly measures enthalpy changes resulting from ligand binding. The binding thermograms for *Bst* $TRAP_{12}$ Δ71 and *Bst* $TRAP_{11}$ could be well fit with single-site (non-cooperative) model, with comparable affinities (Figure 3b; Table S2). However, the binding thermogram for *Bha* $TRAP_{12}$ is biphasic (Figure 3b), indicative of thermodynamically inequivalent binding events. Fitting of the thermograms for *Bha* TRAP to a phenomenological two-site binding model yielded parameters with an initial ($n \approx 1$) ligand-binding event of $K_{d,1} = 1.8 \pm 0.3$ μM, and subsequent ~11 binding events with $K_{d,2} = 0.8 \pm 0.1$ μM. Such parameters would suggest an arguably implausible binding model in which initial weaker binding of the first ligand fundamentally alters the structures of all other binding sites to favor binding, or alternatively, that all subsequent binding events occur in a defined sequential manner. These models describe drastically different modes of cooperativity but are indistinguishable by this method, which provides little insight into the mechanism by which thermodynamic communication occurs between sites.

2.3 | Native MS enables quantification of TRAP states with different numbers of bound ligands

We used native MS to quantify populations of Trp_n-$TRAP_m$ states over a range of Trp concentrations for *Bha* $TRAP_{12}$, *Bst* $TRAP_{12}$ Δ71 and *Bst* $TRAP_{11}$ (Figure 4a). For each protein, titration samples were prepared in parallel by incubating the proteins in 400 mM ammonium acetate solutions with differing concentrations of Trp. For each oligomer, we obtained spectra with narrow charge distributions indicative of native structure.[40] The most abundant charge states for each protein were 24+, 22+ and 21+ for *Bha*, *Bst* Δ71 and *Bst* TRAP, respectively. Recorded spectra provided baseline resolution of signals corresponding to different numbers of bound Trp. Double deconvolution of the *m*/*z* spectra to include proteoforms[37] resulted in very clean zero-charge mass spectra, and improved fits (Figure 4b; Figure S5). Spectra recorded in the absence of Trp contained signals corresponding to the apo oligomeric protein rings, and spectra at the highest Trp concentrations were dominated by the fully bound holo states, while intermediate states could be discerned at all non-zero Trp concentrations. Thus, these spectra enable us to resolve and quantify populations of states corresponding to the ligand-free apo states, the fully liganded holo states, and intermediate species with differing numbers of bound Trp ligands. These populations can be directly compared to the populations predicted by the NN model (Figure 2) and used to determine thermodynamic cooperativity parameters.

2.4 | Differing extent of Trp-Trp cooperativity is evident from population distributions in native MS titrations

Strikingly different population distributions were observed for *Bha*, *Bst* Δ71, and *Bst* TRAP in native MS titrations (Figure 4). For the *Bha* Trp_n-$TRAP_{12}$ titration, distributions were dominated by either Trp_0-$TRAP_{12}$ (apo) or fully bound Trp_{12}-$TRAP_{12}$ (holo) with minimal population of intermediate Trp_n-$TRAP_{12}$ states (Figure 4, left). Such a pattern is consistent with strong positive cooperativity (i.e., Figure 2b). Population distributions for *Bst* TRAP were similar to those previously described,[20] with a high population of intermediate states consistent with weak cooperativity. The distributions for *Bst* TRAP Δ71 reflected an intermediate cooperativity between the *Bst* and *Bha* TRAP (Figure 4, middle). In contrast to the highly congruent CD and ITC titrations, the native MS titration data were quite divergent, implying distinctly different underlying thermodynamics.

Fitting of ligand-dependent population distributions revealed strong cooperativity for *Bha* $TRAP_{12}$, moderate cooperativity for *Bst* $TRAP_{12}$ Δ71, and weak cooperativity for *Bst* $TRAP_{11}$. After correcting binding site concentrations using mass-balance (see Section 5) the experimental Trp_n-$TRAP_{12}$ populations were fit with a binding polynomial corresponding to the NN model. For each dataset, we obtained good agreement between populations determined from integration of deconvoluted spectra, and those simulated from best-fit thermodynamic parameters (Figure 4b, Table 1).

The best fit parameter values obtained from the NN model (Table 1) were consistent with expectations from inspection of the population distribution shifts in the native mass spectra (Figure 4). Cooperativity in *Bst* TRAP yielded a mild eight-fold increase in affinity for sites with one neighbor compared to no neighbors, $\alpha = 8$,

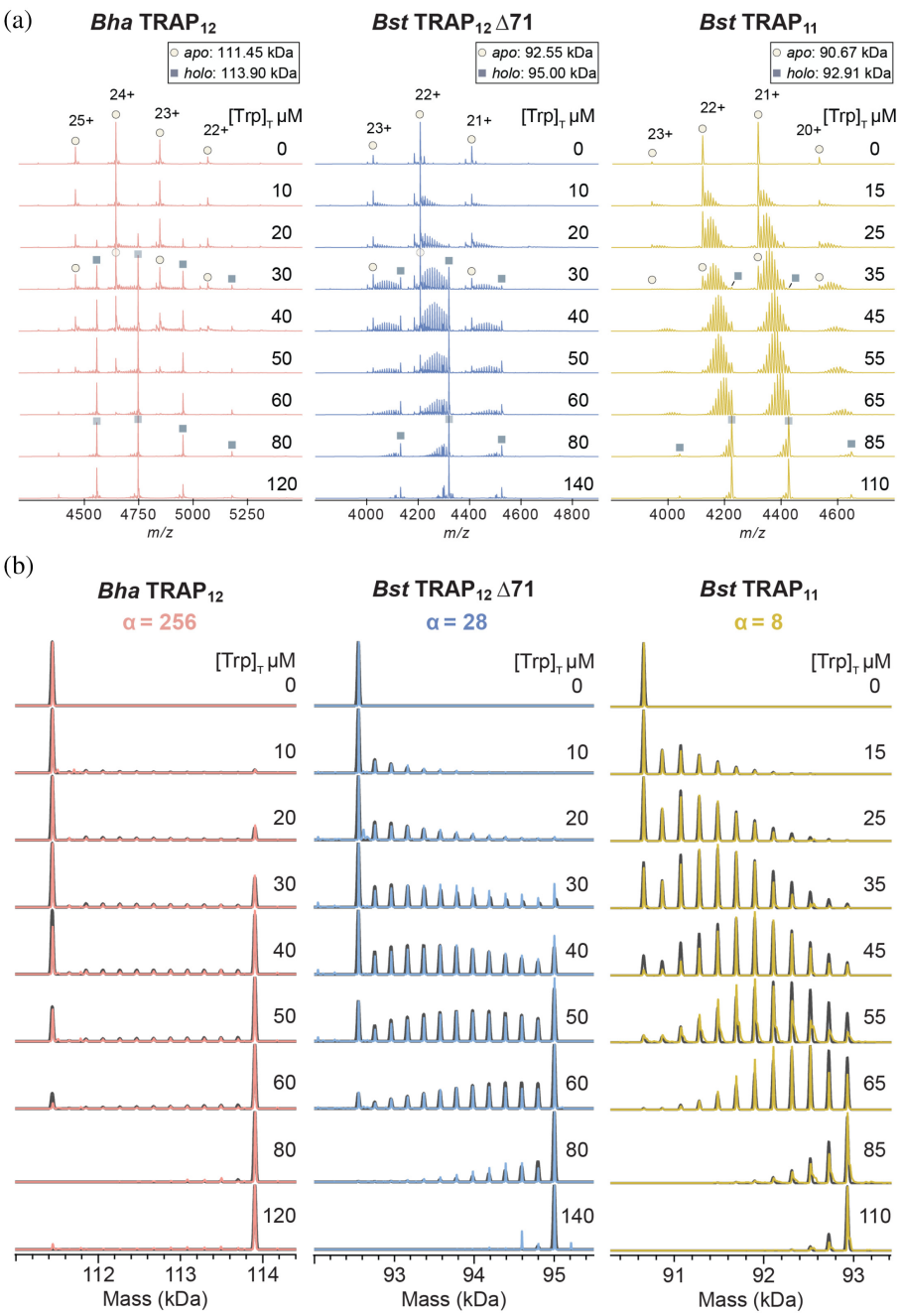

FIGURE 4 Population distributions measured by native MS reveal extents of homotropic cooperativity in ligand binding to 12mer *Bha* and *Bst* Δ71 TRAP (left, middle), and 11mer wild-type *Bst* TRAP (right). (a) Native mass spectra of the three TRAP oligomers (6.0–7.4 μM; Table S3) titrated with 0–140 μM Trp. Increasing Trp in each titration shifts the Trp_n-TRAP populations from apo (Trp_0, circles) to holo ($Trp_{11/12}$, squares). Each Trp titration set with either *Bha* $TRAP_{12}$ (left), *Bst* $TRAP_{12}$ Δ71 (middle), or *Bst* $TRAP_{11}$ (right) features different population distributions, reflecting different degrees of cooperativity. (b) Deconvolved zero-charge mass spectra for the three TRAP oligomers: *Bha* $TRAP_{12}$ (left), *Bst* $TRAP_{12}$ Δ71 (middle), or *Bst* $TRAP_{11}$ (right). Each experimental dataset (colored) is superimposed with a simulated mass spectrum obtained from best-fit parameters to the nearest-neighbor model (grey). Cooperativity is quantified by the parameter α, which is the ratio of the association constant for binding to sites with one bound neighbor, and no bound neighbors (K_{N1}/K_0; Figure 2a). *Bha* $TRAP_{12}$ (left) shows the strongest cooperativity ($\alpha = 256$, $1/K_0 = 718$ μM), with populations dominated by apo and holo rings. Cooperativity is moderate ($\alpha = 28$, $1/K_0 = 217$ μM) in *Bst* $TRAP_{12}$ Δ71 (middle) and weak ($\alpha = 8$, $1/K_0 = 36$ μM) in *Bst* $TRAP_{11}$ (right). TRAP, *trp* RNA-binding attenuation protein

consistent with prior findings.[20] *Bha* TRAP exhibited remarkably strong cooperativity, with a 256-fold higher affinity for sites with one occupied neighbors, consistent with the observation of low abundance for partially loaded states throughout the titration (e.g., Figure 2b). Fits for the engineered 12mer, *Bst* TRAP Δ71 revealed intermediate cooperativity, $\alpha = 28$, consistent with the mixed population distributions of partially and fully loaded states (Figure 4). Thus, the native MS experiments, coupled with statistical thermodynamic modeling enabled an accurate description of the thermodynamics of Trp-TRAP binding for variants with cooperativity values differing by a factor of 32 (=256 ÷ 8).

2.5 | Microscopic parameters predict macroscopic behavior

Typical methods used for measuring bound fractions (e.g., CD, ITC) yield an average signal that is obtained from all extant micro-states and populations, whereas the NN statistical thermodynamic models can define the populations of each individual micro-state. To compare the bulk observable and predicted populations, we simulated macroscopic properties of population distributions of individual protein micro-states over a range of ligand concentrations using the NN model and microscopic parameters obtained from fitting the native MS data

TABLE 1 Best-fit thermodynamic parameters for the NN model of Trp-binding cooperativity for native MS titration data (room temperature ~ 25°C)[a]

	$\Delta G_{N\#}$ (kcal mol^{-1})		$1/K_{N\#}$ (μM)[b]		Cooperativity factor
Protein	ΔG_0	$\Delta G_0 + \Delta G_\alpha$	$1/K_0$	$1/(\alpha K_0)$	α
Bha $TRAP_{12}$	−4.3 ± 0.1	−7.6 ± 0.2	720 ± 140	2.8 ± 0.8	256 ± 87
Bst $TRAP_{12}$ Δ71	−5.00 ± 0.03	−7.0 ± 0.1	217 ± 13	7.8 ± 1.2	28 ± 5
Bst $TRAP_{11}$	−6.1 ± 0.2	−7.3 ± 0.3	36 ± 12	4.4 ± 1.9	8 ± 5

Abbreviations: MS, mass spectrometry; NN, nearest-neighbor; TRAP, *trp* RNA-binding attenuation protein.

[a]Thermodynamic parameters from global fitting of titration data to the nearest-neighbor model, with free energy changes for binding to sites with 0, or 1 occupied neighbors ΔG_{N0}, and $\Delta G_{N1} = \Delta G_0 + \Delta G_\alpha$, respectively. Values are best fit parameters, while uncertainties are standard deviations from three datasets. Standard deviation was calculated between replicates.

[b]Dissociation equilibrium constants are the reciprocal of the corresponding association equilibrium constants.

(Figure 5a,b). These simulations show good agreement between the MS and solution experiments. Discrepancies were largest for the *Bha* TRAP data, whereas they were arguably within experimental error for *Bst* and *Bst* Δ71 TRAP.

Cooperative behavior in ligand binding is often recognized from sigmoidal isotherms that correlate fractional saturation and free ligand concentration. To approximate this behavior, we simulated fractional saturations over various free Trp concentrations [Trp] (Figure 5). The cellular concentration of TRAP protein rings in *B. subtilis* cells has been estimated at ~80 nM (0.9 μM Trp-binding sites),[41] while cellular Trp concentrations have been measured in the low micromolar range.[42] Thus, simulations were performed with a binding site concentration of 1 μM and [Trp] to 20 μM. Compared to experimental titrations which monitor response to total ligand concentration (Figure 5), we see the expected sigmoidal curves predicted for positive cooperativity. Despite 20-fold differences in intrinsic affinities K_0, the apparent affinities fall in a relatively narrow range between ~2 and 8 μM, consistent with those measurements.

We briefly extended the analysis to consider the activation of TRAP for RNA binding. TRAP binds RNA in a Trp-dependent manner, while several binding elements in a single ring are required to achieve high affinity RNA binding.[43,44] To approximate this behavior, we separately compute bulk saturation and the relative abundance of the four basic nearest-neighbor states (empty sites 0, sites with no bound neighbors N0, sites with 1 bound neighbor N1, and sites with two bound neighbors N2) as a function of free ligand concentration [Trp] (Figure 5d). Consistent with the bulk parameters for *Bha* TRAP (Figure 5c), populations of empty sites drop off steeply in exchange for sites with two bound neighbors, whereas for *Bst* and *Bst* Δ71 TRAP states with one bound neighbor become significantly populated. This behavior can be expected to affect the trajectory of Trp-dependent RNA binding.

3 | DISCUSSION

3.1 | Statistical thermodynamic nearest-neighbor models can readily predict complex ligand behavior in cyclic oligomeric proteins

We used a statistical thermodynamic approach to model the binding of ligands to symmetric ring-shaped homo-oligomeric proteins. By applying NN circular lattice model, we generated partition functions corresponding to the 2^N possible liganded states of 12mer and 11mer TRAP rings (Equation (2)). In this model, the Trp-binding sites are arranged on a rigid protein ring, while flexible loops gate access to the binding sites. In their flexible ligand-free states, there is little structural coupling between adjacent sites. Binding of a Trp ligand rigidifies the loops that sandwich its site, resulting in structural, dynamic, and thermodynamic propagation to its adjacent sites (Figure 1d). The model thus enables probabilistic description of a statistically complex ensemble of liganded states using only two microscopic parameters: the intrinsic affinity of a site K_0, and the site–site coupling term α. This parsimonious model allowed us to compute the probability distributions of these states as a function of ligand concentration for a range of cooperativity strengths (Figure 2). Population distributions of states with zero to n bound ligands are revealed to be much more sensitive to the underlying microscopic thermodynamic parameters than the bulk metric of fractional saturation. The ability to measure more directly those underlying population distributions can enable us to decipher mechanisms and magnitudes of cooperativity.

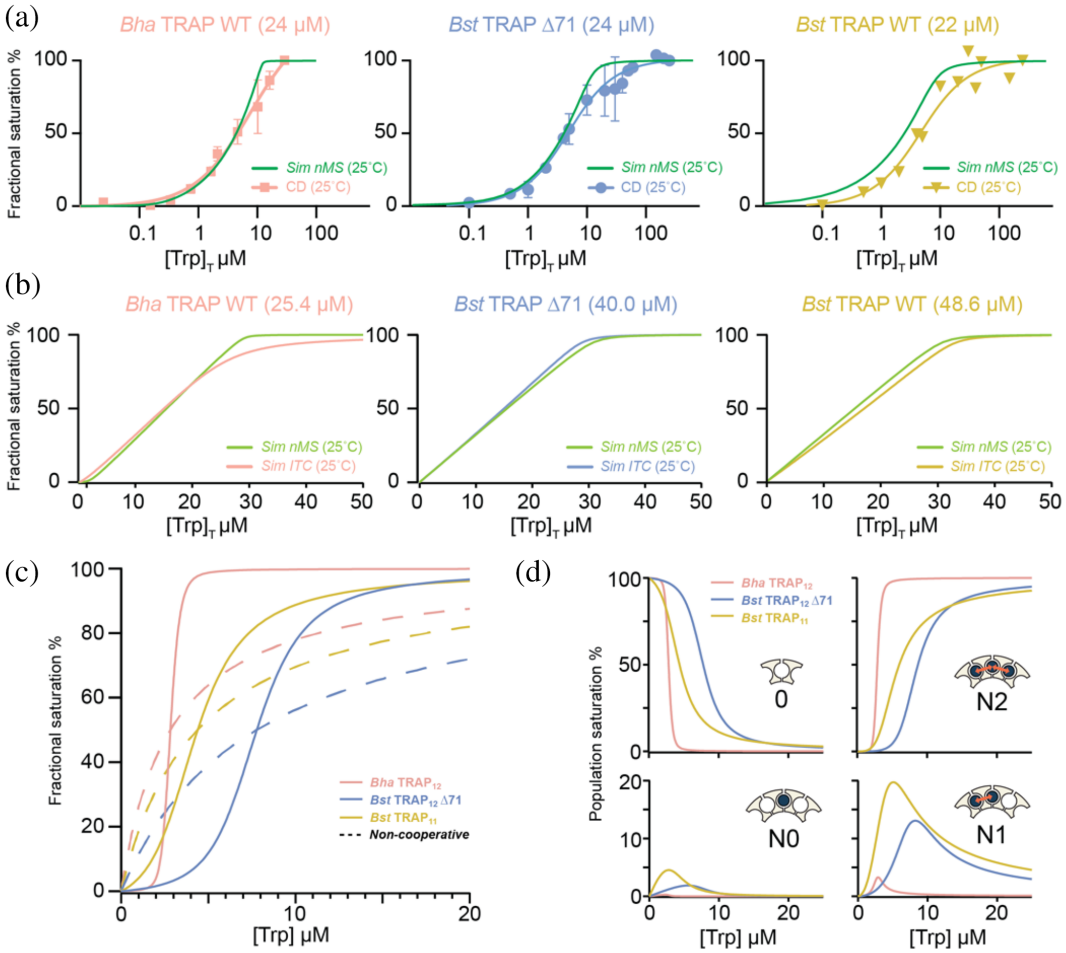

FIGURE 5 Prediction of macroscopic behavior from microscopic parameters from the NN cooperativity model. (a) Simulated native MS bound fraction from nearest-neighbor parameters (green) and experimental CD signal change (at 228 nm) at the indicated Trp-binding site concentrations. (b) Simulated native MS bound fraction nearest-neighbor parameters (green) and computed ITC fractional saturation curves based on phenomenological model parameters ΔG and n (Table S2). Simulations used thermodynamic parameters from native MS at room temperature, and temperature and concentrations used in CD and ITC experiments. (c) Simulated mean site occupancy of TRAP proteins over a range of ligand concentrations [Trp]. *Bha* $TRAP_{12}$ (pink) features a sharp change from almost empty to near complete saturation over the range [Trp] of 2–3 μM. *Bst* $TRAP_{11}$ (yellow) and *Bst* $TRAP_{12}$ Δ71 (blue) have more gradual saturation curves due to their weaker cooperativity. Saturation curves with the same apparent K_D values in the absence of Trp-Trp cooperativity are shown by dashed lines. (d) Population evolution for the four distinct nearest neighbor binding configurations as a function of [Trp]: (1) empty sites (0), (2) bound sites with no bound neighbors (N0), (3) bound sites with one bound neighbor (N1), and (4) bounds sites with two bound neighbors (N2). The strong cooperativity in *Bha* $TRAP_{12}$ (pink) makes N2 the most dominant bound configuration through the range of concentrations. Weaker cooperativity in *Bst* $TRAP_{11}$ (yellow) and *Bst* $TRAP_{12}$ Δ71 (blue) allows significant populations of N0, N1, and N2 states over the concentration range. The concentration of populations with adjacent bound ligands can be understood to affect the protein's regulatory response. CD, circular dichroism; ITC, isothermal titration calorimetry; MS, mass spectrometry; TRAP, *trp* RNA-binding attenuation protein

3.2 | Native MS can directly quantify populations of states with 0 to *n* bound ligands, informing mechanistic binding models

For each of the homo-oligomeric protein variants examined here, native MS enabled us to resolve and quantify populations of oligomers with 0 to n bound ligands (Figure 4). These data showed clear and distinct ligand concentration-dependent population shifts, reminiscent of the population distributions simulated with NN statistical thermodynamic models. Population distributions in each titration series could be fit accurately with binding polynomial derived from this model, yielding thermodynamic parameters that quantify the affinity of a ligand for isolated sites, K_0, and coupling terms that quantify the allosteric thermodynamic communication between sites ΔG_α (Table 1). For *Bst* TRAP this coupling energy was -1.2 ± 0.3 kcal mol^{-1} interaction^{-1}, resulting in an eight-fold higher affinity for binding sites with one

occupied neighbor compared to isolated sites, while for *Bha* TRAP a stronger coupling energy of −3.3 ± 0.2 kcal mol^{-1} $interaction^{-1}$ results in a 256-fold higher affinity. In contrast, conventional ligand-binding measurements by spectroscopic and calorimetric methods (Figure 3) failed to clearly distinguish magnitudes of allosteric coupling. Although the population distributions are strikingly different for the three proteins, bulk properties computed from these distributions were in fact similar, suggesting that the native MS measurements have the unique power to resolve underlying population shifts hidden when averaging over all micro-states.

3.3 | Implications for understanding of TRAP function

TRAP proteins have evolved to regulate expression of the *trp* operon in response to changing levels of free Trp. Although the TRAP protein sequence is highly conserved,[45] disparities exist between proteins from organisms that inhabit vastly different environment conditions (i.e., *Bst*—a thermophile, *Bha*—a halophile), in the encoding of the transacting protein factor anti-TRAP (AT),[46] where the fully transcribed leader possesses sequences capable of preventing ribosome binding to the Shine-Dalgarno sequence,[29] and in their predominant oligomeric states.[47,48] In the three proteins studied here, two 12mers and an 11mer, we observed orders of magnitude differences in Trp-Trp homotropic cooperativity. The origins of these differences merit additional characterization, but it is tempting to speculate that tighter packing of the otherwise isosteric protomers[27] results in stronger coupling between sites. Nevertheless, apparent overall affinities of Trp for *Bha* and *Bst* TRAP were similar, possibly reflecting similarities in the fluctuating concentrations of Trp in bacterial cells.

3.4 | Limitations

The combination of statistical thermodynamic modeling and native MS to extract and quantify allosteric parameters from the partition function is powerful but has limitations. The most obvious is that while NN mechanistic models can be readily envisioned for lattices in which NN interactions can be expected to be dominant, such as symmetric ring-shaped proteins, it is less straightforward to develop such models for hetero-oligomeric proteins, or homo-oligomers with non-cyclic symmetry elements (e.g., double-rings like GroEL[49] and the proteosome[50]). Another limitation has to do with the degeneracy of liganded states: for example, for the NN model, the populations of 12mers with six bound ligands convolve 924 individual configurations with the same mass distributed among six distinct statistical weights (Equation (2)); the abundance of species with six bound ligands does not distinguish between configurations. Thus, it is likely that other models of cooperativity, including MWC-like mechanisms, could provide similar population distributions. Other limitations have to do with the measurements themselves: solution conditions suitable for high resolution native MS typically deviate from the physiological and from solution conditions typical for in vitro biochemical measurements; such deviations could result in context-dependent populations. Moreover, even when native MS-friendly solutions do not perturb populations of states, care must be taken to ensure that ion detection mirrors the populations of states in the solution, avoiding or correcting for differences in ionization efficiencies of different microstates, differences in transmission efficiencies of different ions through the mass spectrometer, or due to potential for dissociation of ligands during ionization and transmission.[22,51,52] The native MS measurements showed that purified TRAP proteins exhibit additional states, including double-rings and numbers of protomers, and possessed proteoforms (Figures S7 and S8). These factors could introduce competing equilibria that complicate efforts to quantify major-state populations in different bound states; of course, such factors would also skew traditional binding measurements that are simultaneously blind to their presence. Nevertheless, with rapidly improving technology for the analysis of proteins by native MS, its combination with statistical thermodynamic models has tremendous potential to advance our understanding of mechanisms of regulation.

4 | CONCLUSIONS

Deciphering mechanisms of allosteric communication between ligand-binding sites in oligomeric proteins is difficult. One serious challenge is that for most biophysical measurements the proportionality between the measured signal and populations of liganded states can be distorted by the very property of interest (i.e., allosteric coupling). As an analytical tool, native MS has the advantage that the signal does not suffer from this confounding effect—the mass unambiguously tells us the number of bound ligands. If the intensity of each signal is proportional to its population, we have an accurate measurement of the underlying distribution of states. Armed with such information, one is then in the position to design and test mechanistic models of site-site communication. In favorable cases, the result can be an understanding of the thermodynamic quantities exchanged between binding sites.

These microscopic thermodynamic parameters provide a genuine basis for establishing the structure–thermodynamic relationships responsible for thermodynamic coupling and biological regulation.

5 | MATERIALS AND METHODS

5.1 | Protein preparation

Three TRAP variants were prepared for these studies: *Bha* TRAP, which forms predominantly 12mers, *Bst* TRAP, which forms predominantly 11mers, and *Bst* TRAP-Δ71, which forms predominantly 12mers.[27,28] Lacking genetically encoded Trp, and bearing only two tyrosines, the proteins share the same low protomer extinction coefficient (ε = 2,980 M^{-1} cm^{-1}) at 280 nm as predicted using ProtParam (https://web.expasy.org/protparam/). Protein concentrations were determined using a Nanodrop 2000c spectrophotometer (Thermo Scientific) in droplet mode.

To produce *Bha* TRAP (WP_010897809.1), the coding region for the wild-type protein, in the pET17b plasmid (Novagen),[29] was modified by the insertion at the C-terminus of sequences encoding the TEV protease cleavage site (ENLYFQ/G) followed by six histidine residues (a His-tag); after removal of the His-tag, the resulting protein consists of ENLYFQ on C-terminus (Figure S2B). Proteins were expressed in *Escherichia coli* BL21(DE3) cells and grown at 37°C in LB. At mid-log phase (OD_{600} = 0.4–0.6), 0.5 mM IPTG (isopropyl β-D-1-thiogalactopyranosid) was added for induction of protein expression. Cells were harvested by centrifugation at 5,000*g* for 5 min, resuspended in lysis buffer (500 mM NaCl, 20 mM NaPi, 100 mM imidazole, pH = 7.4, 1 mM benzamidine, 1 mM PMSF [phenylmethylsulfonyl fluoride]), and lysed in a French pressure cell at 10,000 psi. The lysate was centrifuged at 30,000*g* for 20 min and filtered (0.45 μm cellulose acetate filters, Advantec) before loading onto 1 mL HisTrap™ HP nickel column (GE Healthcare) on an AKTA fast protein liquid chromatography (FPLC) system. His-tagged *Bha* TRAP WT protein was eluted with an imidazole gradient from 100 mM to 1 M in elution buffer (0.1–1 M imidazole, 500 mM NaCl, 20 mM NaPi, pH = 7.4). The pooled protein sample was dialyzed into buffer A (100 mM NaCl, 50 mM $NaPO_4$, pH 8.0) and treated with TEV protease at a mass ratio of 100:1 to cleave the C-terminal His-tag. The protease-treated fraction was then reloaded onto the 1 ml HisTrap™ HP nickel column (GE Healthcare) and eluted with lysis buffer and the flow-through fraction collected. To remove bound Trp, protein was denatured by dialysis for more than 6 hr at room temperature in a 3.5 kDa cut-off dialysis tubing membrane (Spectrum Labs) against 6 M guanidinium hydrochloride (Gdn-HCl) in buffer A. The fraction was then filtered using a 25 mm minisart® syringe filter (pore size: 0.2 μm, Sartorius) and further purified using reversed-phase high-performance liquid chromatography (RP-HPLC) on a 250 × 10 mm Proto 300 C4 column (Higgins Analytical) as previously described.[30] The resulting protein fractions were lyophilized and stored at −80°C until needed. The protein was refolded by dissolving the lyophilized pellet in 6 M Gdn-HCl buffer A at a concentration of ∼1 mg ml^{-1}, and gradually refolded by sequential dialysis against 3 M, 1.5 M, 0.75 M, and 0 M Gdn-HCl in buffer A in 3.5 kDa cut-off dialysis membrane (Spectrum Labs) for 12 hr per round at room temperature. The refolded protein was then purified by size-exclusion chromatography using a Hiload™ 16/600 Superdex™ 75 pg column (Sigma-Aldrich) in buffer A.

The gene of *Bst* TRAP WT (WP_033013997.1) and the C-terminal-five-residue-truncated *Bst* TRAP Δ71 were expressed from pET17b and pET9a vectors (Novagen, Inc.), respectively, as previously described.[20,27] The proteins were separately expressed in *E. coli* BL21(DE3) cells, grown at 37°C in LB, and induced at OD_{600} = 0.4–0.6 with 0.5 mM IPTG. Cells were harvested by centrifugation, resuspended in lysis buffer B (100 mM K_2HPO_4, 50 mM KCl, 1 mM EDTA, pH = 8.0) plus 1 mM benzamidine and 1 mM PMSF (phenylmethylsulfonyl fluoride). The cells were broken in a French pressure cell at 10,000 psi. The lysate was clarified by centrifugation at 30,000*g* for 25 min. For every 5 ml of supernatant, 1 ml of 2% protamine sulfate was added. Then the lysate was stirred on ice for 30 min. The lysate was centrifuged at 30,000*g* for 25 min and the supernatant was heated to 80°C for 10 min followed by another centrifugation at 30,000*g* for 25 min. The supernatant was then dialyzed overnight against 50 mM Tris–HCl (pH = 8.0). Dialyzed proteins were loaded on Mono Q™ 10/100 GL column (GE Healthcare) in 50 mM Tris–HCl (pH 8.0). Proteins were eluted with a salt gradient from 0 to 1 M NaCl in 50 mM Tris–HCl (pH 8.0). The same protocol and buffers described above were used to remove Trp and refold proteins to obtain apo (ligand-free) proteins, except that the 0.75 M Gdn-HCl dialysis step was omitted during refolding, and that RP-HPLC purified *Bst* TRAP Δ71 was refolded at 55°C instead of at room temperature.

5.2 | Circular dichroism

Tryptophan binding to TRAP was measured by CD spectroscopy at 25°C using a JASCO model J-715 spectropolarimeter with cell path length of 1 mm. Spectra were

recorded between 226 nm and 230 nm using ~24 μM TRAP protomers (~2 μM ring) in 50 mM sodium phosphate buffer, pH 8.0, in the presence of increasing concentrations of Trp (0–300 μM). The change in CD signal at 228 nm, which shows the maximal difference between spectra of apo- and Trp-liganded TRAP, was used as a measure of tryptophan binding.[31] The change in CD signal was fit with the quadratic single-site binding equation (Equation (1)).

5.3 | Isothermal titration calorimetry

RP-HPLC-purified, refolded, and SEC-purified TRAP proteins were dialyzed extensively against buffer A. A stock solution of tryptophan was prepared by dissolving L-Trp powder (Sigma Aldrich, 93,659) in the dialysate. Trp concentration was determined from its extinction coefficient at 280 nm of $\varepsilon = 5{,}540\ M^{-1}\ cm^{-1}$ using a Nanodrop 2000c spectrophotometer (Thermo Scientific) in droplet mode. The measured concentrations of Trp in the syringe and of TRAP protomers in the cell were 293 μM and 23 μM for *Bha* TRAP, 433 μM and 38 μM for *Bst* TRAP Δ71, and 645 μM and 74.7 μM for *Bst* TRAP WT, respectively. ITC thermograms were recorded on a MicroCal VP-ITC (Malvern Instruments) at 25°C. A reference power of 30 μcal s^{-1} and a stirring speed of 307 rpm were applied for all experiments. Deionized water was used in the reference cell. The software NITPIC was used to correct the baseline for each isotherm by iteratively optimizing the noise parameter *wrmsd* and then integrating the enthalpy at each titration point.[32] The integrated heats showed bimodal binding behavior for *Bha* TRAP, and single mode for *Bst* Δ71 and *Bst* TRAP. The integrated enthalpy data were then fit with two-sites or one-site binding models using *itcsimlib*.[5] TRAP oligomer concentrations were corrected assuming a stoichiometry of 1 Trp per protomer based on the fitted stoichiometry n; corrected concentrations were 25.4 μM for *Bha* TRAP, 40.0 μM for *Bst* TRAP Δ71, and 48.6 μM for *Bst* TRAP WT. Confidence intervals in fitted parameters were estimated from 200 bootstrapped samples.

5.4 | Native mass spectrometry

Aliquots of purified proteins were dialyzed into 400 mM ammonium acetate (Sigma Aldrich, 431,311). For *Bha* TRAP and *Bst* TRAP Δ71, the pH was adjusted to pH 8 by adding ammonium hydroxide (~1 M NH_3 in H_2O, Sigma-Aldrich). Dialyzed protein concentrations (~60 μM protomers) were measured as above ($\varepsilon = 2{,}980\ M^{-1}\ cm^{-1}$) at 280 nm using a Nanodrop 2000c spectrophotometer (Thermo Scientific) in droplet mode. The Trp was dissolved in 400 mM ammonium acetate and the Trp concentration was measured as above ($\varepsilon = 5{,}540\ M^{-1}\ cm^{-1}$) at 280 nm.[33] Samples for each titration point were prepared freshly in parallel by mixing the dialyzed protein solutions with serial dilutions from a 1 M Trp stock solution to the indicated ligand concentrations (0–140 μM).

Native MS experiments were performed at room temperature (~25°C) on a Q Exactive Ultra-High Mass Range (UHMR) Orbitrap mass spectrometer (Thermo Fisher Scientific) modified to allow for surface induced dissociation, similar to that previously described.[34] Emitters were pulled from borosilicate filament capillaries (OD 1.0 mm, ID 0.78 mm, Sutter Instrument) in-house on a P-97 Flaming/Brown Micropipette Puller (Sutter Instrument) that is fitted with a thermometer placed near the heating element. The temperature near the heating element was stabilized below 30°C between each pull, and the same tip pulling parameters were used for each titration set. Each sample was loaded into an emitter using gel loading tips (Genesee Scientific, 0.5-10 μl Ultra Micro Gel Tip) to avoid contamination, and then placed on a Nanospray Flex ion source (Thermo Fisher Scientific) equipped with a platinum wire to allow for static nano electrospray ionization. To obtain a stable total ion current, the spray voltage was initially ramped to about 1.4 kV and then gradually decreased to between 0.6–0.9 kV and held constant. For all measurements, the following instrument tune settings were kept constant: capillary temperature 250°C, Source DC Offset 21 V, S-lens RF level 200, detector optimization to low *m/z*, ion transfer target high *m/z*, injection flatapole DC 5 V, inter flatapole lens 4 V, bent flatapole DC 2 V, transfer multipole DC 0 V, C-trap entrance lens inject 1.8 V, HCD energy 1 V, HCD field gradient 200 V, and HCD cell pressure 4 (UHV Sensor ~3−4E-10 mbar). For most emitters, spectra were collected at resolutions of 6,250, 12,500 and 25,000, as defined at 400 *m/z*. Each spectrum was obtained by averaging the same number of scans, and the injection time and averaged micro-scans were fixed. For *Bha* TRAP and *Bst* TRAP Δ71, each titration point was recorded three times using different emitters. For *Bst* TRAP, two sets of spectra were obtained from different time points from a single emitter, and the third sample from a different emitter.

Data were manually examined using Xcalibur Qual Browser software (Thermo Fisher Scientific), and spectral deconvolution was carried out using UniDec.[35] For quantification of Trp_n-$TRAP_m$ populations, all spectra from each titration set were initially deconvolved using MetaUniDec V4.4 and then later each spectrum was reprocessed using UniDec V5.0.1 to allow for the double deconvolution approach to be used to produce deconvolved zero-charge mass spectra where the proteoforms

found in the apo TRAP spectrum from each titration are combined.[36,37] The titration spectra collected at a resolution of 12,500 were used for the final deconvolutions. Custom deconvolution parameters were used for each protein titration set. For all sets, UniDec parameters were adjusted to a narrow m/z range around the dominant oligomer charge state distribution, the charge range was set to less than 11 charges surrounding the most abundant charge state; mass sampling was set every 1 Da, FWHM (full width at half maximum) was set to 0.8 Th, peak shape function was set to split Gaussian/Lorentzian; artifact suppression, point smooth width and mass smooth width were set to 0; charge smooth width was set to 1.5, native charge offset range was set to ±6, and m/z to Mass transform was set to interpolate. A Mass List Window width of ±1,200 Da was used. A mass list for each protein was generated based on the protein mass, oligomeric state N, and 0 to N bound tryptophan. For *Bst* TRAP and *Bst* TRAP Δ71 this list consisted of the expected 11mer and 12mer oligomeric states, respectively, with variable Trp. For *Bha* TRAP, we found that in addition to the expected 12mer, minor overlapping signals from trace amounts of a 13mer were present, so the mass list was set to include those species. The respective deconvolved apo TRAP zero-charge mass spectrum was used as the double deconvolution kernel file. The areas for each species in the deconvolved spectra were extracted, sum normalized, and used for the model fitting. The area extraction window was set to ±2 times the standard deviation. The standard deviation was determined based on the average FWHM of the deconvolved apo and holo peaks using the simplified equation where FWHM divided by 2.35 equals the standard deviation.

Accurate determination of protein concentration is complicated for TRAP proteins because they lack encoded Trp residues and have low extinction coefficients, while a 5–10% error in concentration translates to ±1 site for an 11–12 mer protein. Before thermodynamic analysis of the MS titration data, protein concentrations were corrected by fitting the bound fraction at each Trp concentration with the quadratic independent sites model; $[L]_T$ was assumed to be accurate, while $[M]_T$ and the apparent dissociation constant K_D were fit parameters.[20] This fit yielded the binding site concentration $[M]_T$ that was in general about 33% higher than initial estimated concentration (Table S3, Figure S6).

5.5 | Fitting of population distributions with the NN model

Following TRAP concentration correction, experimental data obtained from native MS were fit using a NN model for 12 binding sites via *itcsimlib*[5], as described previously.[20] Briefly, the relative abundance of Trp-TRAP species (0–12 bound, 13 total liganded states) was determined from the integrated peak areas; because Trp binding in excess of the available sites is not observed, we did not consider the possibility of non-specific binding of Trp to TRAP. To fit this population distribution, the NN model was parametrized with values for ΔG_0 (intrinsic binding free energy) and ΔG_α (binding-free energy with an occupied neighbor). A binding partition function that includes all possible Trp_n-$TRAP_{12}$ configurations ($2^{12} = 4{,}096$) was generated using *itcsimlib*. The statistical thermodynamic probabilities of each configuration were then computed at each experimental ligand and binding site concentration. The probabilities of each of 4,096 configurations were summed and converted into the relative abundance of the 13 Trp-TRAP species, permitting generation of a virtual titration dataset. Then the discrepancy between native MS experimental titration data and the virtual titration data was globally minimized by iteratively varying the parameters of ΔG_0 and ΔG_α using a Powell[38] gradient-descent optimization algorithm. Uncertainties in fit parameters were obtained from the standard deviation of three independent experimental repeats.

5.6 | Population simulations with NN models

Simulated population distributions for 12-mer proteins were computed using NN statistical thermodynamic models, with the same $\Delta G_{N2} = -7.5$ kcal mol^{-1} (binding-free energy with two occupied neighbors) but varied NN cooperativity factor α, and therefore different ΔG_0 (intrinsic binding free energy).[5] In the NN model, given ΔG_0 and ΔG_α values, the statistical thermodynamic probability of each of Trp_n-$TRAP_{12}$ configuration ($2^{12} = 4{,}096$), along with the probabilities of the four basic NN energy configurations (0, N0, N1, N2) were computed over a range of 280 ligand concentrations using *itcsimlib*.[5] From these trajectories we computed fractional saturation curves by summing the three ligand-bound states, N0, N1, and N2, as a function of free ligand concentration [L] (Figure 2c; Figure S3). The relative abundance of species with 0–12 bound Trp was used to simulate Gaussian peak shapes for each state (Figure 2b). To obtain apparent Hill coefficients, the resulting trajectories were fit using the Hill equation, $Y = \frac{[L]^h}{[L]^h + K_{d,app}^h}$, where h is the Hill coefficient and $K_{d,app}$ is the apparent equilibrium dissociation constant.

AUTHOR CONTRIBUTIONS

Weicheng Li: Conceptualization (equal); formal analysis (equal); methodology (equal); software (equal); visualization

(equal); writing – original draft (equal); writing – review and editing (equal). **Andrew S. Norris:** Conceptualization (equal); data curation (equal); formal analysis (equal); investigation (equal); methodology (equal); validation (equal); visualization (equal); writing – review and editing (equal). **Katie Lichtenthal:** Conceptualization (equal); data curation (equal); investigation (equal); resources (equal); validation (equal). **Skyler Kelly:** Data curation (equal); investigation (equal); resources (equal). **Elihu C. Ihms:** Conceptualization (supporting); methodology (supporting); software (equal). **Paul Gollnick:** Conceptualization (equal); project administration (equal); resources (lead); supervision (equal). **Mark P. Foster:** Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); software (equal); supervision (equal); validation (equal); visualization (equal); writing – review and editing (equal).

ACKNOWLEDGMENTS

We thank Dr. Melody "Pepsi" Holmquist for help with native MS experiments. We also thank members of the Foster lab, Gollnick Lab, and Wysocki lab for useful discussions.

FUNDING INFORMATION

This work was supported by National Institutes of Health (NIH) grants R01 GM120923 (to Mark P. Foster and Paul Gollnick), R01 GM077234 (to Mark P. Foster and Paul Gollnick), and P41 GM128577 (to Vicki H. Wysocki).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Weicheng Li https://orcid.org/0000-0001-6356-2674
Andrew S. Norris https://orcid.org/0000-0002-0121-5922
Katie Lichtenthal https://orcid.org/0000-0002-1717-3114
Skyler Kelly https://orcid.org/0000-0002-6841-0044
Elihu C. Ihms https://orcid.org/0000-0003-3034-1685
Paul Gollnick https://orcid.org/0000-0003-3293-1363
Vicki H. Wysocki https://orcid.org/0000-0003-0495-2538
Mark P. Foster https://orcid.org/0000-0001-9645-7491

REFERENCES

1. Goodsell DS, Olson AJ. Structural symmetry and protein function. Annu Rev Biophys Biomol Struct. 2000;29:105–153.
2. André I, Strauss CEM, Kaplan DB, Bradley P, Baker D. Emergence of symmetry in homooligomeric biological assemblies. Proc Natl Acad Sci U S A. 2008;105:16148–16152.
3. Stefan MI, Le Novère N. Cooperative binding. PLoS Comput Biol. 2013;9:e1003106.
4. Brown A. Analysis of cooperativity by isothermal titration calorimetry. Int J Mol Sci. 2009;10:3457–3477.
5. Ihms EC, Kleckner IR, Gollnick P, Foster MP. Mechanistic models fit to variable temperature calorimetric data provide insights into cooperativity. Biophys J. 2017;112:1328–1338.
6. Cui Q, Karplus M. Allostery and cooperativity revisited. Protein Sci. 2008;17:1295–1307.
7. Dyachenko A, Gruber R, Shimon L, Horovitz A, Sharon M. Allosteric mechanisms can be distinguished using structural mass spectrometry. Proc Natl Acad Sci U S A. 2013;110:7235–7239.
8. Walker TE, Shirzadeh M, Sun HM, et al. Temperature regulates stability, ligand binding (Mg2+ and ATP), and stoichiometry of GroEL-GroES complexes. J Am Chem Soc. 2022;144:2667–2678.
9. Hill A, The V. Combinations of Haemoglobin with oxygen and with carbon monoxide. Int Biochem J. 1913;7:471–480.
10. Nishikori S, Esaki M, Yamanaka K, Sugimoto S, Ogura T. Positive cooperativity of the p97 AAA ATPase is critical for essential functions. J Biol Chem. 2011;286:15815–15820.
11. Huang R, Ripstein ZA, Rubinstein JL, Kay LE. Cooperative subunit dynamics modulate p97 function. Proc Natl Acad Sci U S A. 2019;116:158–167.
12. Hilser VJ, Wrabl JO, Motlagh HN. Structural and energetic basis of allostery. Annu Rev Biophys. 2012;41:585–609.
13. Gollnick P, Babitzke P, Antson A, Yanofsky C. Complexity in regulation of tryptophan biosynthesis in Bacillus subtilis. Annu Rev Genet. 2005;39:47–68.
14. Otridge J, Gollnick P. MtrB from Bacillus subtilis binds specifically to trp leader RNA in a tryptophan-dependent manner. Proc Natl Acad Sci U S A. 1993;90:128–132.
15. Antson AA, Dodson EJ, Dodson G, Greaves RB, Chen XP, Gollnick P. Structure of the trp RNA-binding attenuation protein, TRAP, bound to RNA. Nature. 1999;401:235–242.
16. McElroy C, Manfredo A, Wendt A, Gollnick P, Foster M. TROSY-NMR studies of the 91 kDa TRAP protein reveal allosteric control of a gene regulatory protein by ligand-altered flexibility. J Mol Biol. 2002;323:463–473.
17. Kleckner IR, McElroy CA, Kuzmic P, Gollnick P, Foster MP. Homotropic cooperativity from the activation pathway of the allosteric ligand-responsive regulatory trp RNA-binding attenuation protein. Biochemistry. 2013;52:8855–8865.
18. Pauling L. The oxygen equilibrium of hemoglobin and its structural interpretation. Proc Natl Acad Sci. 1935;21:186–191.
19. Saroff HA, Kiefer JE. Analysis of the binding of ligands to large numbers of sites: The binding of tryptophan to the 11 sites of the trp RNA-binding attenuation protein. Anal Biochem. 1997;247:138–142.
20. Holmquist ML, Ihms EC, Gollnick P, Wysocki VH, Foster MP. Population distributions from native mass spectrometry titrations reveal nearest-neighbor cooperativity in the ring-shaped oligomeric protein TRAP. Biochemistry. 2020;59:2518–2527.
21. Bennett JL, Nguyen GTH, Donald WA. Protein-small molecule interactions in native mass spectrometry. Chem Rev. 2021;122:7327–7385. https://doi.org/10.1021/acs.chemrev.1c00293.

22. Karch KR, Snyder DT, Harvey SR, Wysocki VH. Native mass spectrometry: Recent progress and remaining challenges. Annu Rev Biophys. 2022;51:157–179.
23. Holt JM, Ackers GK. Chapter 7 The Hill coefficient: Inadequate resolution of cooperativity in human hemoglobin. Methods Enzymol. 2009;455:193–212.
24. Bucci E, Pucciarelli S, Angeletti M. Free energy changes and components implicit in the MWC allosteric model for the cooperative oxygen binding of hemoglobin. Biochemistry. 2013;52: 4149–4156.
25. Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: A plausible model. J Mol Biol. 1965;12:88–118.
26. Koshland DE, Nemethy JG, Filmer D. Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry. 1966;5:365–385.
27. Chen CS, Smits C, Dodson GG, et al. How to change the oligomeric state of a circular protein assembly: Switch from 11-subunit to 12-subunit TRAP suggests a general mechanism. PLoS One. 2011;6:e25296.
28. Chen XP, Antson AA, Yang M, et al. Regulatory features of the trp operon and the crystal structure of the trp RNA-binding attenuation protein from Bacillus stearothermophilus. J Mol Biol. 1999;289:1003–1016.
29. Szigeti R, Milescu M, Gollnick P. Regulation of the tryptophan biosynthetic genes in bacillus halodurans: Common elements but different strategies than those used by Bacillus subtilis. J Bacteriol. 2004;186:818–828.
30. McElroy CA, Manfredo A, Gollnick P, Foster MP. Thermodynamics of tryptophan-mediated activation of the trp RNA-binding attenuation protein. Biochemistry. 2006;45:7844–7853.
31. Li PTX, Gollnick P. Characterization of a trp RNA-binding attenuation protein (TRAP) mutant with tryptophan independent RNA binding activity. J Mol Biol. 2004;335:707–722.
32. Scheuermann TH, Brautigam CA. High-precision, automated integration of multiple isothermal titration calorimetric thermograms: New features of NITPIC. Methods. 2015; 76:87–98.
33. Mach H, Middaugh CR, Lewis RV. Statistical determination of the average values of the extinction coefficients of tryptophan and tyrosine in native proteins. Anal Biochem. 1992;200:74–80.
34. Vanaernum ZL, Gilbert JD, Belov ME, Makarov AA, Horning SR, Wysocki VH. Surface-induced dissociation of noncovalent protein complexes in an extended mass range Orbitrap mass spectrometer. Anal Chem. 2019;91:3611–3618.
35. Marty MT, Baldwin AJ, Marklund EG, Hochberg GKA, Benesch JLP, Robinson CV. Bayesian deconvolution of mass and ion mobility spectra: From binary interactions to polydisperse ensembles. Anal Chem. 2015;87:4370–4376.
36. Reid DJ, Diesing JM, Miller MA, et al. MetaUniDec: High-throughput deconvolution of native mass spectra. J Am Soc Mass Spectrom. 2019;30:118–127.
37. Norris CE, Keener JE, Perera SMDC, et al. Native mass spectrometry reveals the simultaneous binding of lipids and zinc to rhodopsin. Int J Mass Spectrom. 2021;460:116477.
38. Fletcher R, Powell MJD. A rapidly convergent descent method for minimization. Comput J. 1963;6:163–168.
39. Ben-Naim A. Cooperativity and regulation in biochemical processes. New York, NY: Springer Science & Business Media, 2001. https://doi.org/10.1007/978-1-4757-3302-0.
40. Kafader JO, Melani RD, Schachner LF, et al. Native vs denatured: An in depth investigation of charge state and isotope distributions. J Am Soc Mass Spectrom. 2020;31:574–581.
41. McCabe BC, Gollnick P. Cellular levels of trp RNA-binding attenuation protein in Bacillus subtilis. J Bacteriol. 2004;186: 5157–5159.
42. Røst LM, Brekke Thorfinnsdottir L, Kumar K, et al. Absolute quantification of the central carbon metabolome in eight commonly applied prokaryotic and eukaryotic model systems. Metabolites. 2020;10:74.
43. Li PTX, Scott DJ, Gollnick P. Creating hetero-11-mers composed of wild-type and mutant subunits to study RNA binding to TRAP. J Biol Chem. 2002;277:11838–11844.
44. Li PTX, Gollnick P. Using hetero-11-mers composed of wild type and mutant subunits to study tryptophan binding to TRAP and its role in activating RNA binding. J Biol Chem. 2002;277: 35567–35573.
45. McAdams NM, Patterson A, Gollnick P. Identification of a residue (Glu60) in TRAP required for inducing efficient transcription termination at the trp attenuator independent of binding tryptophan and RNA. J Bacteriol. 2017;199:e00710-16.
46. Valbuzzi A, Yanofsky C. Inhibition of the B. subtilis regulatory protein TRAP by the TRAP-inhibitory protein, AT. Science. 2001;293:2057–2059.
47. Ruotolo BT, Giles K, Campuzano I, Sandercock AM, Bateman RH, Robinson CV. Biochemistry: Evidence for macromolecular protein rings in the absence of bulk water. Science. 2005;310:1658–1661.
48. Watanabe M, Heddle JG, Kikuchi K, et al. The nature of the TRAP-anti-TRAP complex. Proc Natl Acad Sci U S A. 2009;106: 2176–2181.
49. Saibil HR, Fenton WA, Clare DK, Horwich AL. Structure and allostery of the chaperonin GroEL. J Mol Biol. 2013;425:1476–1487.
50. Huang X, Luan B, Wu J, Shi Y. An atomic structure of the human 26S proteasome. Nat Struct Mol Biol. 2016;23:778–785.
51. Page JS, Kelly RT, Tang K, Smith RD. Ionization and transmission efficiency in an electrospray ionization-mass spectrometry interface. J Am Soc Mass Spectrom. 2007;18:1582–1590.
52. Rovelli G, Jacobs MI, Willis MD, Rapf RJ, Prophet AM, Wilson KR. A critical analysis of electrospray techniques for the determination of accelerated rates and mechanisms of chemical reactions in droplets. Chem Sci. 2020;11:13026–13043.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Li W, Norris AS, Lichtenthal K, Kelly S, Ihms EC, Gollnick P, et al. Thermodynamic coupling between neighboring binding sites in homo-oligomeric ligand sensing proteins from mass resolved ligand-dependent population distributions. Protein Science. 2022; 31(10):e4424. https://doi.org/10.1002/pro.4424