Donut dynamics: localized allosteric activation and ligand-binding kinetics in the ring-shaped ligand-responsive regulatory protein TRAP

Ian R Kleckner (Biophysics Program)

Advisor: Mark P Foster (Biochemistry)

The Ohio State University

2010 Hayes Graduate Research Forum

May, 2010 in Columbus, OH

First place, Math and Physical Sciences division

Abstract

The overarching goal of this research is to gain insight into mechanisms of ligand-modulated gene regulation by studying the *trp* RNA-binding Attenuation Protein (TRAP). TRAP is a biological Trp-sensor endogenous in bacilli: when cellular free Trp concentration is low, TRAP predominates in the inactive apo form and the micro-organism may synthesize additional Trp using a specific molecule of RNA. However, when the Trp concentration rises TRAP may bind this excess Trp and thus become activated to bind that RNA molecule. When TRAP binds the RNA, it disables downstream production of additional Trp. Thus, the micro-organism employs TRAP in this simple feedback mechanism to ensure that it will only synthesize Trp if it is scarce.

The goal of the current study is to rigorously and quantitatively examine the details of TRAP activation which occur upon Trp-binding using two distinct approaches:

(1) NMR spectroscopy is used to probe the structure and dynamics of TRAP. This study quantifies μs-ms molecular dynamics present in apo TRAP which are hypothetically required for

proper function. Upon Trp-binding, there are quantifiable changes in both the structure and dynamics of TRAP. Importantly, the protein appears to rigidify at the μs-ms timescale.

(2) Stopped-flow fluorescence spectroscopy is used to investigate the time course of Trp binding upon rapid mixing with apo TRAP. These data reveal not one, but two kinetic steps interpreted as (a) Trp-binding which occurs slower than the diffusion limited rate, followed by (b) conformation exchange of TRAP between two distinct states. These data also indicate formation of bonds, increase in order and compression of TRAP structure upon binding.

These investigations help Biologists to understand TRAP at a mechanistic level. They are also of broad interest towards rigorously quantifying the time-dependent behavior of any ligand-responsive protein (many of which are essential for life).

Introduction

Gene regulation is central to all living organisms and can occur at many hierarchical levels in cells including DNA replication, RNA transcription/maturation and protein translation. Many disease states, including cancer, are a direct result of specific failures of these regulatory mechanisms. Proteins are the principal endogenous molecular machines that carry out gene regulation. The proteins act by binding other partner molecules (ligands) to modulate activity and achieve coordinated regulation. Much like a protein's structure, their dynamic motions (or lack thereof) are responsible for proper function and are indirectly encoded by the amino acid sequence [1]. The importance of regulatory RNA and protein dynamics are evident even in the simple organism *B. subtilis*, which has long-served as an excellent model for studying gene regulation [2]. These basic principles motivate the current *in vitro* study to gain insight into mechanisms of ligand-mediated gene regulation by the *trp* RNA-binding Attenuation Protein (TRAP).

Some microorganisms can synthesize the essential amino acid tryptophan via genes of the *trp* operon. In some bacilli, this process is regulated by the protein TRAP via a simple feedback mechanism. TRAP forms a radially symmetric donut-shaped 11-mer with binding sites for 11 Trp molecules. In the absence of Trp (apo), TRAP is inactive for binding RNA and therefore the transcription of *trp*-mRNA proceeds unattenuated (figure S1, solid line). In contrast, when Trp is abundant, TRAP will bind up to 11 Trp molecules (holo) and become activated for *trp*-mRNA-binding. This binding prohibits further transcription of the *trp*-mRNA and subsequent expression of the *trp* genes (figure S1, dashed line). This mechanism ensures that the organism will only express the proteins that synthesize Trp when the amino acid is scarce.

The current study seeks to address the mechanism of TRAP activation after tryptophan binding. Prior studies support a hypothesis in which apo TRAP is flexible at the μs-ms timescale and the RNA binding site is therefore poorly structured. Upon binding Trp, the structure of TRAP is perturbed and rigidified at this timescale to permit high-affinity RNA binding (K_d ~ nM) [3, 4, 5]. This hypothesis motivates several questions investigated here: what changes in the structure and dynamics of TRAP occur upon tryptophan binding? How are these changes related to TRAP activation? What is the role of μs-ms dynamics in the function of TRAP? Is the rate of tryptophan binding gated by TRAP? What is the rate of TRAP activation after tryptophan binding?

First, TRAP structure and dynamics are examined in the free vs. bound state and observed dynamic motions are quantitatively described. This is accomplished using NMR spectroscopy, custom software development and statistical analysis of the data. Subsequently, the kinetics of Trp-binding are studied to determine the rate binding and to gain structural insight into these time-dependent states. This is accomplished using stopped-flow fluorescence spectroscopy, statistical model selection and classical thermodynamic analyses.

Results and Discussion

Tryptophan binding alters the structure of TRAP: NMR chemical shift perturbations

The TRAP protein studied here refers to *B. stearothermophilus* A26I TRAP unless indicated otherwise. TRAP contains 7 Ile (labeled with ^{1}H and ^{13}C at δ_1), 4 Leu (labeled with ^{1}H and ^{13}C at both δ_1 and δ_2) and 7 Val (labeled with ^{1}H and ^{13}C at both γ_1 and γ_2) yielding a total of 29 methyl probes. ^{1}H and ^{13}C stereo-assignments of all 29 methyl groups were completed for both apo and holo TRAP using 3D spectroscopy of ^{1}H, ^{13}C and ^{15}N nuclei (see materials and

methods). The 29 labeled methyl groups permit observation of the core, subunit-subunit interfaces and one of the tryptophan binding loops (BC loop).

^{1}H-^{13}C-HSQC 2D correlation spectra of both apo and holo TRAP are shown in figure 1. The position of each peak in the spectrum reports on the local chemical environment of that methyl group, commonly interpreted as an abstract metric for local structure [6]. Upon binding tryptophan, each peak is perturbed to a new position, thus reporting its new chemical environment. The difference in peak position between the apo and holo states is referred to as the chemical shift perturbation (CSP) and its magnitude is a site-specific metric for the structural change induced upon tryptophan binding.

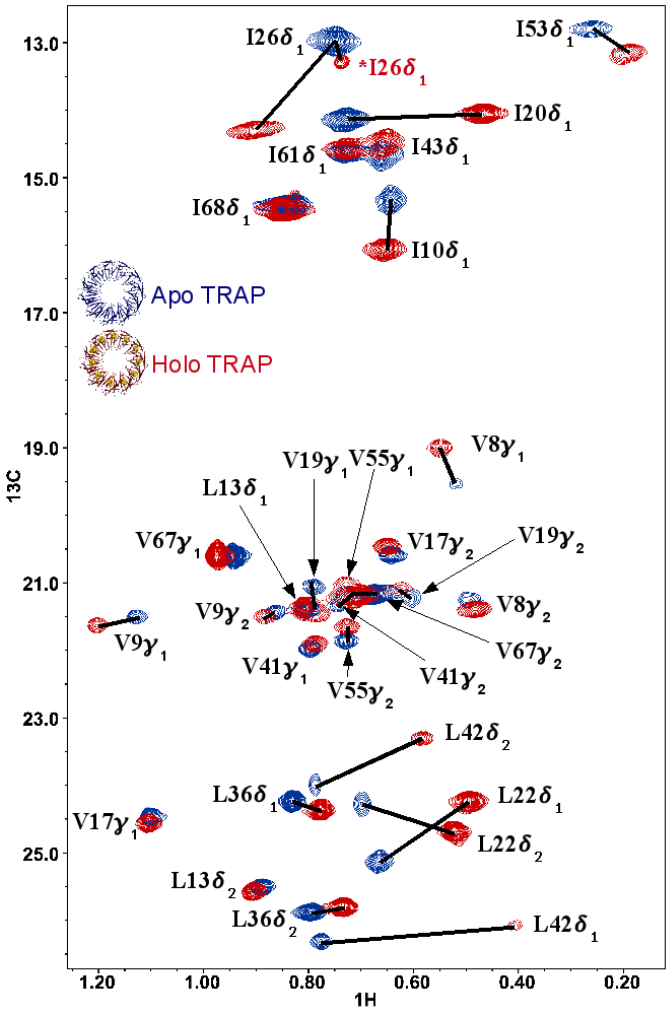


Figure 1. Tryptophan binding induces site-specific changes in the structure of TRAP. Overlay of Trp-free (apo, blue) and Trp-bound (holo, red) TRAP ^{1}H-^{13}C-HSQC 2D correlation spectra at 25°C. Black lines connect the apo and holo crosspeaks for the 29 methyl probes. The minor holo peak for Ile 26δ_1 at 13.5 ppm ^{13}C (marked *) indicates this methyl group populates both a holo state and a minor apo-like state.

The CSPs are illustrated by the black lines connecting apo and holo crosspeaks in figure 1 and are mapped to a model of *B. stearothermophilus* holo WT TRAP bound to RNA 1C9S [5] in figure 2a. This striking result indicates that the largest CSPs upon addition of tryptophan occur near the tryptophan binding BC loop and RNA binding site. The aromatic indole ring on the side chain of tryptophan produces a small local magnetic field which can alter the chemical shift of nuclei within ~5.5 Å. The program SHIFTS [7] was used to estimate the effect of the bound

tryptophan's ring current shift and confirms the CPSs observed here are in large part due to relevant structural changes of TRAP (i.e., not just the presence of tryptophan's ring current). This structural change upon binding tryptophan makes sense mechanistically since apo TRAP is unable to bind RNA.

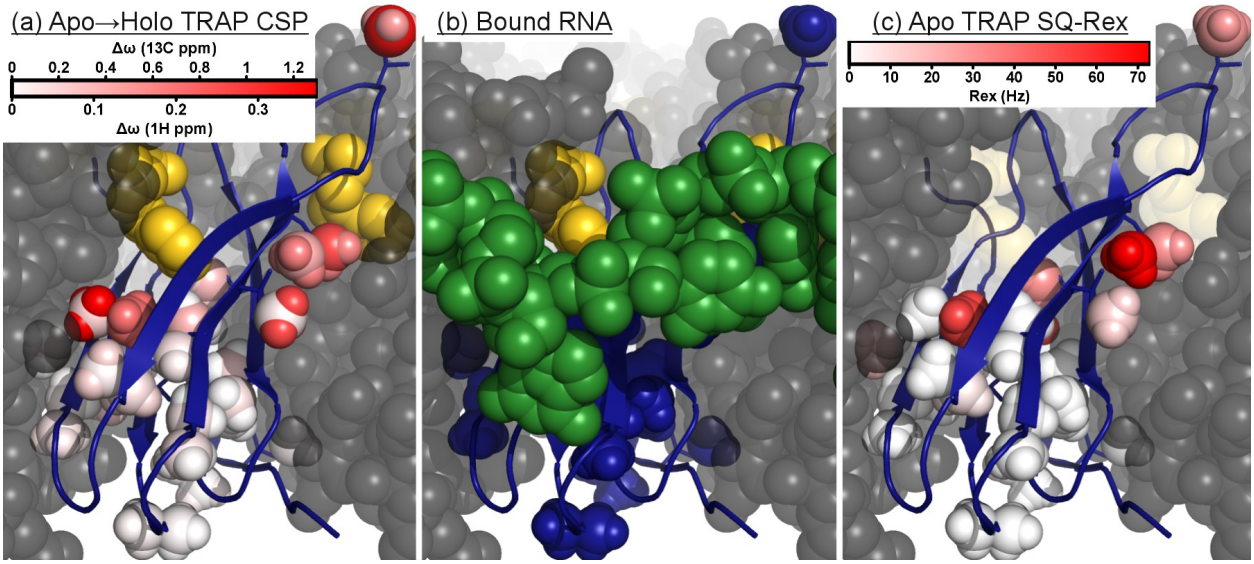


Figure 2. Localized changes in structure and dynamics of TRAP upon tryptophan binding. Chemical shift perturbations (CSPs) and methyl relaxation dispersion (MRD) data are mapped to the crystal structure 1C9S of holo WT TRAP bound to RNA. Single TRAP protomer in blue cartoon, neighboring protomers in gray spheres, bound Trp as gold spheres, 29 methyl probes as colored spheres, RNA as green spheres in b only. **(a)** ^{1}H and ^{13}C CSPs between apo and holo TRAP indicate local structural changes and are limited to the TRAP core, interface and Trp-binding loop. **(b)** The RNA-binding site coincides with largest apo-holo CSPs and apo R_{ex}. **(c)** Exchange broadening R_{ex} of apo TRAP reports on μs-ms flexibility which appears to be limited to the TRAP core, interface and Trp-binding loop

The CSP for Ile 26δ1 in the BC loop indicates a structural change which may be involved in accommodating the ligand's entry. In the Trp-bound state, this methyl group apparently occupies one of two distinct local structures denoted holo-like and apo-like based on chemical shift (apo-like state marked "*" in figure 1). Interestingly, the methyl groups distant from the RNA-binding surface do not undergo large chemical shift changes (e.g., at bottom of figure 2a).

This supports the hypothesis that the structural changes induced upon tryptophan binding are limited to a regulatory "hot spot" which activates TRAP for binding *trp*-mRNA.

Typtophan binding eliminates flexibility of apo TRAP: NMR methyl relaxation dispersion (MRD)

To investigate the dynamics of TRAP at the μs-ms timescale, ^{13}C single-quantum methyl relaxation dispersion (^{13}C-SQ MRD) was preformed on both apo and holo TRAP [8]. In NMR, if a structural probe yielding an observable signal flexible at the μs-ms timescale, the relaxation rate R_2^{eff} may be increased because the probe spends time in multiple states during the detection period of the experiment. This phenomenon is manifested as a reduction in signal intensity and an increase in signal linewidth (broadening). This "chemical exchange broadening" can be partially eliminated by applying a series of equally-spaced 180° radio-frequency (RF) pulses. In the experiments performed here, ^{13}C RF pulses are applied during a defined relaxation delay. This so-called CPMG pulse train will refocus the observed signal broadening due to μs-ms dynamics. The result is typically presented for each signal as a "dispersion curve" which plots the effective relaxation rate R_2^{eff} as a function of the CPMG pulse train frequency.

Figure 3a (data in blue) illustrates the characteristic decay of $R_2^{eff}(\nu_{CPMG})$ measured at 10 CPMG frequencies for two methyl groups in apo TRAP. As ν_{CPMG} increases, the exchange-broadening and therefore relaxation rate R_2^{eff} is reduced (hence "relaxation dispersion"). A two-state exchange model described below was fit to these data to quantify the amount of exchange broadening $R_{ex} = R_2^{eff}(0) - R_2^{eff}(\infty)$. R_{ex} is a simple and attractive metric for quantifying μs-ms dynamics since its value is relatively insensitive to the accuracy of the parameters in the fitted model of exchange ($|\Delta\omega_C|$, P_A, k_{ex} and R_2^0). R_{ex} values measured using ^{13}C-SQ MRD at 25°C for the 29 probes in apo TRAP are mapped to the structure in figure 2c. The methyls within 10-15 Å

of the RNA-binding surface exhibit significant exchange broadening. The associated μs-ms flexibility may indirectly mask the RNA-binding surface of apo TRAP to prevent the RNA from binding. Coincidental with the CSPs in figure 2a, the methyl groups distant from the RNA-binding surface do not exhibit large R_{ex} in apo TRAP (e.g., base of the protein).

Upon binding tryptophan, the μs-ms dynamics in the methyl groups of apo TRAP are abolished. This can be seen in figure 3a, where for Ile $10\delta_1$ R_{ex} is reduced from 35.8 Hz to 0 Hz and for Ile $26\delta_1$ R_{ex} is reduced from 29.9 Hz to 0. The F-test was used at 1% to determine the presence of a significant dispersion (i.e., if $R_{ex} > 0$). These results indicate that tryptophan binding induces localized changes in both structure and dynamics of TRAP. These changes may be mechanistically required to permit the RNA to bind.

Describing apo TRAP dynamics: rigorous MRD analysis

The goal of this section is to more rigorously describe the dynamic motions observed in apo TRAP. The R_{ex} values shown in figure 2 are informative but only serve as a starting point for investigating the rates and magnitudes of motion. To this end, additional MRD data were acquired and a custom Matlab software package was developed to fit a model of protein motion. In this model, each methyl group is assumed to independently undergo exchange between two distinct local structures denoted A and B. This may physically correspond to a minor adjustment of the side chain or perhaps a more significant global unfolding process. Each site-specific minor state (B) is lowly populated (2-15% here) and is not directly observed in the NMR spectrum. However, the presence of each B state is *indirectly* evidenced by exchange-broadening of the NMR peak for the directly observed A state (i.e., $R_{ex} > 0$). Note that the NMR signals in figure 1 correspond to these directly observed A states for each of the 29 structural probes. The A↔B

exchange is described by the chemical shift difference between the states for ^{1}H and ^{13}C ($|\Delta\omega_H|$, $|\Delta\omega_C|$), the intrinsic relaxation rate of the average state (R_2^0) and a pair of kinetic terms (either k_A and k_B *or* k_{ex} and P_A). k_A and k_B are the rates of conversion from A→B and B→A respectively, $k_{ex} = k_A+k_B$ is the total exchange rate and $P_A = k_B / (k_A+k_B)$ is the population-fraction of the A state.

Proper analysis of MRD data requires a significant amount of independent observations to balance the large number of fitting parameters [9, 10]. This was accomplished here by performing additional experiments were under selected conditions: ^{13}C-SQ and ^{1}H-^{13}C-MQ [10] experiments at 3 temperatures (25, 32 and 37°C) using 10 distinct CPMG frequencies at 800 and 600 MHz ^{1}H static fields to produce a total of 172 2D NMR spectra. From these spectra, 7-10 dispersion curves were extracted per NMR signal (e.g., figure 3b). This yields 70-100 independent observations with 18 independent fitting parameters: $|\Delta\omega_H| = |\omega_H^A - \omega_H^B|$, $|\Delta\omega_C| = |\omega_C^A - \omega_C^B|$, $P_A(25°C)$, $k_{ex}(25°C)$, $P_A(32°C)$, $k_{ex}(32°C)$, $P_A(37°C)$, $k_{ex}(37°C)$, and 7-10 R_2^0 values (one for each dispersion curve). To ensure fitting accuracy, a 4D grid search was performed to sample initial fitting conditions and a map of χ^2 space was used to seek and/or exclude alternate fitting solutions (see materials and methods).

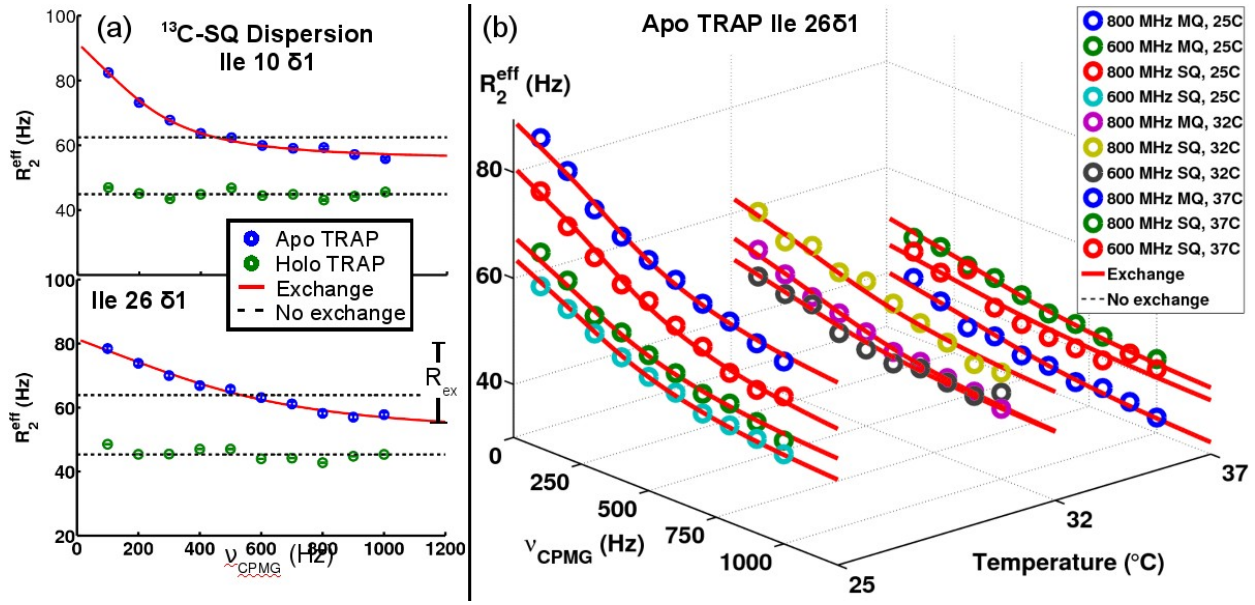


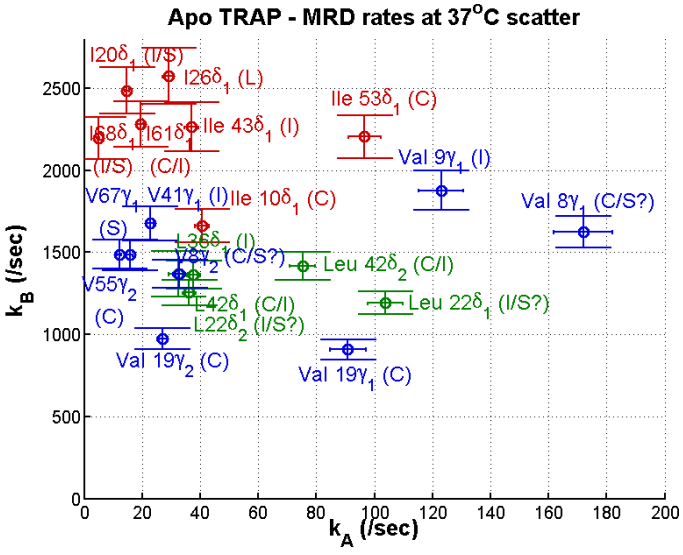
Figure 3. Tryptophan binding eliminates the μs-ms flexibility present in apo TRAP evidenced by methyl relaxation dispersion (MRD) **(a)** Medium to large dispersions for TRAP residues Ile 10δ_1 and Ile 26δ_1 obtained from ^{13}C-SQ methyl relaxation dispersion at 25°C. The exchange broadening $R_{ex} = R_2^{eff}(0) - R_2^{eff}(\infty)$ is a simple metric for μs-ms dynamics. Red lines represent a statistically significant fit to the two-state exchange model and dashed lines represent a fit to a model of no-exchange. Since R_{ex} is reduced significantly upon addition of tryptophan, it appears the protein is rigidified at this timescale. **(b)** Additional MRD data acquired to describe the motions of apo TRAP. Each crosspeak of apo TRAP is independently fit to the two-state exchange model A↔B using the 7-10 dispersion curves acquired for that crosspeak (only Ile 26δ1 shown here). This analysis permits accurate description of the rates and magnitudes of motion in apo TRAP. The data are fit such that each of 29 crosspeaks has chemical shift differences $|\Delta\omega_H|$, $|\Delta\omega_C|$, each of 3 temperatures (for each crosspeak) has equilibrium population P_A and exchange rate $k_{ex} = k_A + k_B$, and each of 10 dispersion curves (for each crosspeak) has relaxation rate R_2^0.

Out of the 29 methyl groups in apo TRAP, 22 groups favored the exchange model (i.e., $R_{ex} > 0$) and 20 groups yielded reasonable estimates for all the fitted parameters (Val 9γ_2 and Val 55γ_1 had ill-defined P_A resulting from fast exchange $k_{ex} >> |\Delta\omega|$). In what follows, the parameters $|\Delta\omega_H|$, $|\Delta\omega_C|$, and the temperature-dependence of $k_A = P_B k_{ex}$ and $k_B = P_A k_{ex} = k_{ex} - k_A$ are quantitatively interpreted to form a model of the motions of apo TRAP. Importantly, since there is no high resolution structural model for apo TRAP, the crystal structure 1C9S of holo TRAP is used as a guide [5]. This work indicates that holo TRAP is structurally and dynamically *different*

than apo TRAP. Nevertheless, the holo TRAP structure is an essential starting point for hypothesizing a high-resolution structural and dynamical description of the apo protein.

Motions within apo TRAP are asynchronous: MRD k_A,k_B analysis

In this model, each methyl group is assumed to *independently* undergo exchange between two distinct local structures denoted A and B. However, it is possible that some methyl groups are subject to concerted/synchronized motion. If a set of methyl groups are reporting exchange from the same dynamic process then there should be agreement in the forward and reverse rates of exchange (k_A and k_B) as well as the temperature-dependence of those rates (quantified here by ΔH, ΔS and $E_A(A \rightarrow B)$ and $E_A(B \rightarrow A)$). That is, their motions would remain synchronized at all temperatures (like a pair of windshield wipers on a car that remain synchronized regardless of the wiper-speed setting).



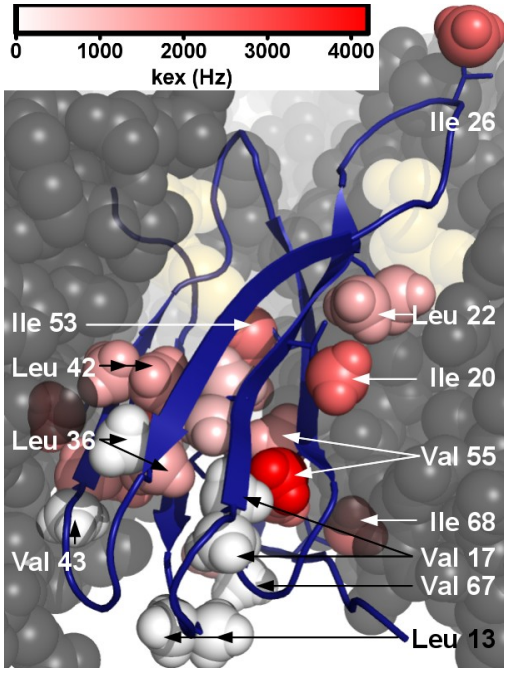


Figure 4. Kinetic rates of site-specific A↔B exchange in apo TRAP reflect independent motions due to the large variance in (k_A,k_B) values. **Left:** The (k_A,k_B) plot illustrates partitioning of dynamic modes based on side chain identity (Ile in red, Leu in green, Val in blue) and structural location in holo TRAP (C=Core, I=Interface, L=Loop, S=Solvent). **Right**: The total rate of exchange $k_{ex}=k_A+k_B$ is mapped to the structure of holo TRAP and reveals a flexible core and interface at 1-2 kHz and more flexible surface at ~2.5 kHz with a rigid base. Data shown are at 37ºC to illustrate maximum differences in site-specific rates.

Figure 4 illustrates the (k_A,k_B) plot at 37ºC for the 20 exchanging probes. Due to the large variance in these site-specific values, there are necessarily asynchronous/independent motions within the protein. That is, some methyl groups will be in their A state while some will be in their B state. The state of apo TRAP at any point in time can be considered a combination of A and B states for each of these independent motions in TRAP. Since holo TRAP exhibits no such dynamics, the state of its methyl probes appear to be locked into a single configuration at the µs-ms timescale.

Upon closer inspection of figure 4a, there appears to be kinetic partitioning based on side chain identity (Ile vs. Leu vs. Val) and structural location. The Ile residues are shown in red and occupy a limited region of the (k_A,k_B) plot with exception of I10 and I53 which are both buried in the core. The Leu and Val residues exhibit a larger range of exchange kinetics which have not yet been interpreted in terms of the structure. Interestingly, Leu and Val side chains each have two methyl probes (δ_1 & δ_2 and γ_1 & γ_2 respectively) but despite their structural proximity on the same side chain, these methyl pairs sometimes report apparently different dynamic motions. For L36, V41 and V67, only one methyl probe exhibits $R_{ex} > 0$ and for both V8 and V9, the pair of methyl probes are distant in the (k_A,k_B) plot. This may occur if the side chain is fixed in the structure and rate of change of the surrounding environment is different for the two methyl probes. Perhaps one methyl is pointed toward the protein core while the other is pointed toward

solvent. This may differentially affect the rigidity of the Cδ-Cγ bond (Leu) or the Cγ-Cβ bond (Val) for the two methyls on the side chain. Torsion about this bond vector is a common mechanism for changes in chemical shift of the terminal methyl carbon [6]. Importantly, NMR spectra in figure 1 indicate that the two methyl probes have different chemical shifts reporting different time-averaged local structure (e.g., δ_1 and δ_2 NMR peaks are distant for L13, L42). Therefore it seems reasonable that they may report different time-dependent local structure as well. Of course this dynamic discrepancy could also be explained by a lack of fitting accuracy, but this does not appear to be the case based on χ^2 maps used to assess fitting accuracy (not shown). Therefore, this affect is believed to be real.

Figure 4 shows the k_{ex} values mapped to the structure of TRAP and reveals a more complete motional description of the apo protein compared to simply using the R_{ex} values in figure 1c. It appears that the residues buried in the core (V19, I53) and interface (L22, L42, I68, V41) exhibit exchange around k_{ex} = 1000, 2300 and 1290, 1400, 2200, 1700 Hz. The side chains on the surface (I20, I26, I68) exchange more quickly with k_{ex} = 2500, 2600, 2200 Hz. This indicates the buried residues exchange around 1-2 kHz with Ile faster than Leu and Val and the surface residues exchange more quickly around 2.5 kHz. The residues near the base of the protein exhibit $R_{ex} = 0$ which may simply be a result of continuous solvent-exposure which does not warrant the chemical shift difference required for detection via the NMR RD experiment (i.e., if $\Delta\omega = 0$ then $R_{ex} = 0$ and the structural probe appears to be rigid).

In order to quantify the thermodynamics of the exchange event at each probe, the van't Hoff analysis is performed (see materials and methods). This quantifies the temperature-dependence of the equilibrium constants $K = k_A/k_B$ to extract site-specific enthalpy and entropy changes ΔH and ΔS upon exchange from A to B. Each of the methyl groups probed exhibit a

negative change in both enthalpy and entropy upon exchange to their B state, consistent with a more-bonded and more-ordered B state. The results are shown in figure 5 and much like the (k_A,k_B) plot there appears to be partitioning based on side chain identity (Ile vs. Leu vs. Val) and to a lesser extent on structural location.

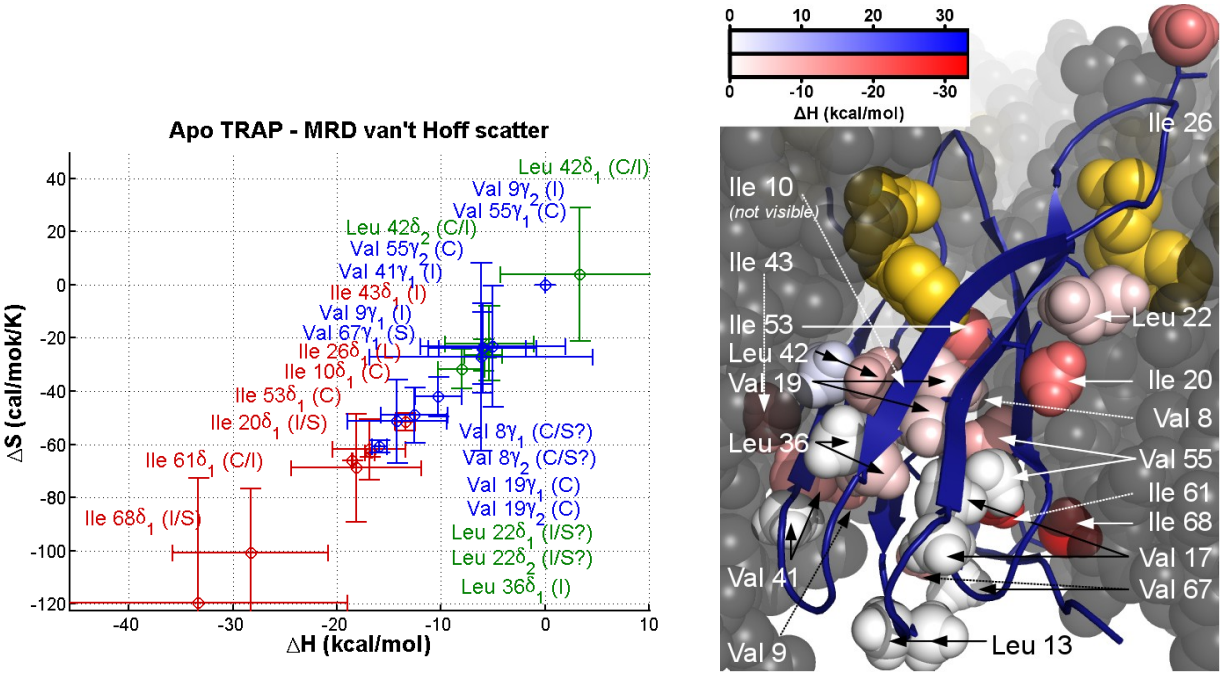


Figure 5. The dynamics of apo TRAP exhibit energetic partitioning based on side chain identity and structural location. **Left**: van't Hoff scatter plot of site-specific change in enthalpy ΔH and entropy ΔS upon exchange from A→B. These values appear to be partitioned by side chain identity (Ile, Leu or Val) and to a lesser extent structural location (C=Core, I=Interface, S=Solvent, L=Loop). Ile residues exhibit the largest changes in enthalpy and entropy whereas Leu and Val residues are more similar. **Right:** ΔH values are mapped to the structure of TRAP indicate the observed side-chain-type partitioning is more significant than the observed structure-based partitioning.

This energetic partitioning may reflect more general feature of side chain structure rather than simply a special feature of the TRAP protein. This hypothesis will be pursued in future studies. Arrhenius analyses of the temperature-dependence of the rates k_A and k_B provide activation energies of the A↔B exchange. Unfortunately interpretation was unclear due to large

fitting errors which rendered the site-specific probes as indistinguishable by this metric (data not shown).

Apo TRAP samples structures unlike the holo state: MRD $|\Delta\omega_H|$ and $|\Delta\omega_C|$ analysis

To investigate the magnitude and nature of the motions of apo TRAP, the site-specific chemical shift differences |Δω| from the RD analysis are examined. Each |Δω| value reflects the chemical shift difference between the two local structures for the NMR probe.

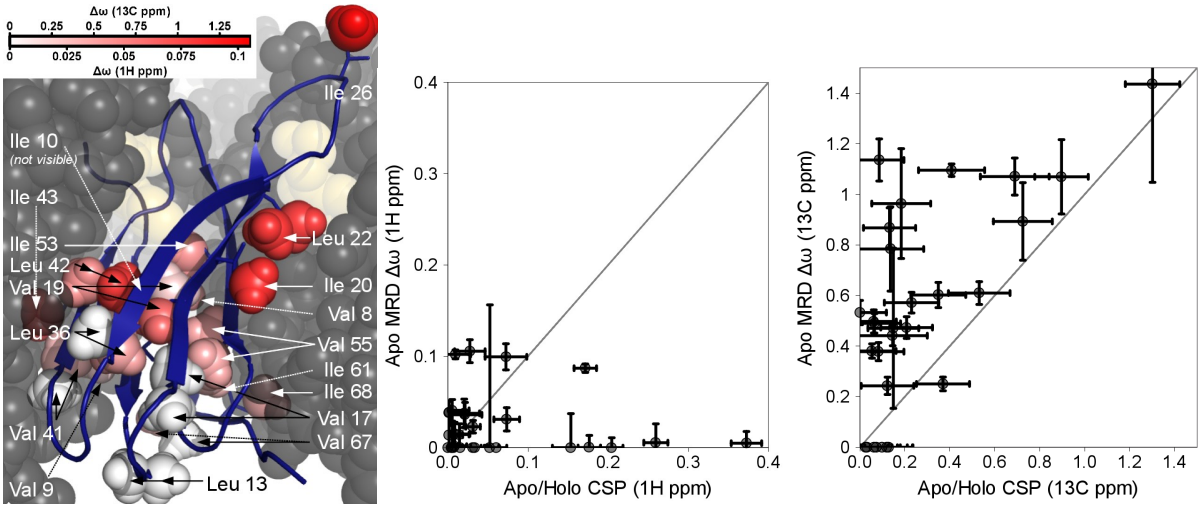


Figure 6. Analysis of apo TRAP dynamics indicate the largest local structure changes occur at the protein surface and that the structures sampled are unlike the Trp-bound state. **Left:** Chemical shift differences Δω from fits to RD data indicate the largest changes in structure occur at the protein surface. Hypothetically this reflects exchange between a buried local structure and solvent-exposed local structure. **Center** (and **Right)**: ^{1}H (and ^{13}C) chemical shift differences Δω from RD are compared to chemical shift changes Δδ upon addition of tryptophan for each site-specific probe. If apo TRAP samples local structures similar to the Trp-bound state, there will be agreement between these values (i.e., points will lie along the diagonal). However, the majority of outlying data points indicate that apo TRAP samples unique local structures unlike the Trp-bound state.

Figure 6, left indicates that the largest chemical shift differences upon site-specific exchange occur at the surface of TRAP (e.g,. Ile 26, Leu 22, Ile 20, Leu 42). Hypothetically,

these side chains are exchanging between a buried local structure and a solvent-exposed local structure as they flip into and out of the protein. The probes in the core of the protein report a smaller and more uniform chemical shift difference which likely reflects only a minor re-adjustment between the two local structures. Side chains at the base of the protein (Leu 13, Val 67) report zero chemical shift difference which, as explained above, may reflect constant exposure to solvent and/or structural rigidity.

To test the hypothesis that the motions of apo TRAP sample a holo-like state, these site-specific RD parameters $|\Delta\omega_H|$ and $|\Delta\omega_C|$ are compared to the apo-holo CPSs $\Delta\delta_H$ and $\Delta\delta_C$ measured from the NMR spectra in figure 1. If the probes in apo TRAP sample local structures similar to their Trp-bound state, there will be agreement between these values and points will lie along the diagonal of figure 6, center and right. The many non-diagonal points indicate that apo TRAP samples unique local structures unlike the Trp-bound state.

TRAP binds tryptophan binds in two kinetic steps: stopped-flow fluorescence

To investigate the mechanism of TRAP activation, stopped-flow kinetics experiments were performed by monitoring the time course of the tryptophan ligand's fluorescence upon rapid-mixing with TRAP. The fluorescence increases when tryptophan enters the hydrophobic binding pocket of TRAP and time-resolved data reveal two kinetic steps which are completed in ~1 sec and 10-30 sec respectively (figure 7).

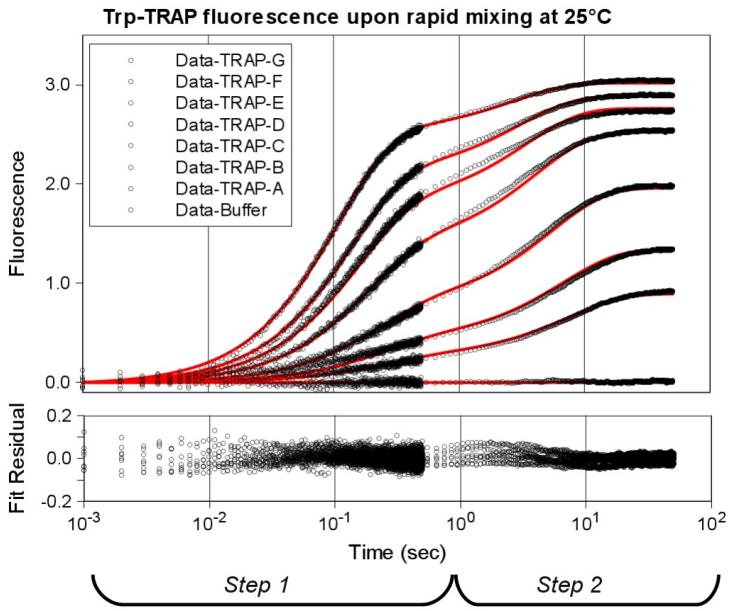


Figure 7. Time-resolved tryptophan binding experiments reveal a multi-step binding mechanism. Tryptophan fluorescence transients were acquired upon rapidly mixing 1 μM tryptophan with 1-22 μM TRAP (A through G). The data are fit to a two-step model describing bi-molecular binding followed by uni-molecular TRAP isomerization: mechanism (M1).

Fluorescence transients were acquired using 1 μM tryptophan mixed with 1-22 μM TRAP (monomer concentration) at seven temperatures (15, 20, 25, 30, 35, 40, 45ºC). Data at each temperature were globally analyzed using the program Dynafit [11] which enabled rigorous analysis of 10 hypothetical binding mechanisms (see appendix). The analysis indicates that the most reasonable mechanism to model the data describes a bi-molecular binding step followed by a uni-molecular TRAP isomerization step.

$$Trp + TRAP \underset{k1B}{\overset{k1F}{\Leftrightarrow}} TrpTRAP \underset{k2B}{\overset{k2F}{\Leftrightarrow}} TrpTRAP^{*} \quad \text{(M1)}$$

Table 1. Fit parameters from analysis of kinetic data using (M1)

Temp (ºC)	k1F (/M/sec)	k1B (/sec)	k2F (/sec)	k2B (/sec)	R(TW)	R(TW*)
15	247043 ± 1003	0.853 ± 0.007	0.172 ± 0.002	0.028 ± 0.000	0.716 ± 0.002	0.611 ± 0.001
20	362136 ± 1164	1.665 ± 0.010	0.317 ± 0.003	0.051 ± 0.000	0.757 ± 0.002	0.608 ± 0.001
25	537453 ± 2009	3.039 ± 0.018	0.440 ± 0.004	0.097 ± 0.000	0.766 ± 0.003	0.616 ± 0.001
30	768851 ± 3233	5.324 ± 0.028	0.701 ± 0.006	0.176 ± 0.001	0.720 ± 0.003	0.581 ± 0.001
35	1005540 ± 4854	10.083 ± 0.048	1.155 ± 0.011	0.382 ± 0.002	0.714 ± 0.003	0.547 ± 0.001
40	1198400 ± 12380	21.189 ± 0.148	2.029 ± 0.031	0.959 ± 0.006	0.804 ± 0.008	0.535 ± 0.002
45	1239700 ± 28730	42.186 ± 0.345	3.179 ± 0.093	1.721 ± 0.008	0.943 ± 0.022	0.485 ± 0.003

It is important to acknowledge that the accuracy of the parameters which describe some physical model are limited to the accuracy of the model employed [12]. In this case, it may be more physically accurate to describe this kinetic system by an additional binding step with another requisite isomerization:

$$Trp + TRAP \underset{k1B}{\overset{k1F}{\Leftrightarrow}} TrpTRAP \underset{k2B}{\overset{k2F}{\Leftrightarrow}} TrpTRAP^{*} + Trp \underset{k3B}{\overset{k3F}{\Leftrightarrow}} Trp_2TRAP^{*} \underset{k4B}{\overset{k4F}{\Leftrightarrow}} Trp_2TRAP^{**} \quad \text{(M2)}$$

This introduces an additional 6 parameters to the model (4 rates and 2 quantum yields). Presumably, the isomerization event in (M1) encompasses any additional binding and isomerization steps not explicitly described by the mechanism. That is, k_2 from mechanism (M1) is some composite of k_2, k_3 and k_4 from mechanism (M2). Since even the three-step models tested were too complex for these data, mechanism (M1) appears to be the best approach in light of the simplifications it affords. Statistical consideration of 9 alternate mechanisms is discussed in the appendix. Furthermore, it is worth acknowledging that the accuracy of spectrophotometric-based analyses (e.g., fluorescence) has been under careful scrutiny considering discrepancies with calorimetric measurements [13, 14, 15, 16, 17]. This has been attributed to an inaccurate

physical mechanism and/or limitations of the van't Hoff and Arrhenius analyses [12]. However, these data decidedly inform on the requirement of multiple kinetic steps, the timescale of initial binding, the timescale of the subsequent events and the energetic partitioning between these steps.

Tryptophan binding is retarded 10^4x: interpreting the first step

The goal of this section is to formulate a model describing the encounter of tryptophan and TRAP in solution. The rate constant for diffusion-limited encounter between Trp and the TRAP binding site have been examined with three theoretical models discussed by Shoup, et al. [18]. The three models serve to make an accurate prediction of the encounter rate and provide reasonable upper and lower limits for comparison.

$$\text{Upper Limit} \quad k_{\text{Smol}} = 4\pi N_0 \frac{100\, R_{\text{Enc}} \left(D_{\text{Trp}} + D_{TRAP} \right)}{1000}$$

$$\text{Lower limit} \quad k_{\text{SS}} = k_{\text{Smol}} \left(\frac{A}{4\pi R_{Enc}^2} \right)$$

$$\text{Most accurate} \quad k_{\text{Shoup}} = k_{\text{Smol}} \left(\frac{\sqrt{(A/\pi)}}{\pi R_{Enc}} \right)$$

N_0 is Avagadro's number, R_{Enc} is the distance of closest approach in meters, D is the diffusion coefficient in cm^2/sec, A is the total area of binding sites on the TRAP donut in $meters^2$ and the factors of 100 and 1/1000 yield rate constants in /M/sec. The upper limit k_{Smol} is provided by the Smoluchowski equation and estimates the encounter rate constant between a spherical TRAP of radius R_{Enc} and the small ligand tryptophan. The second model yields a lower limit by

accounting for the fractional area of TRAP which is eligible for tryptophan binding. The rate constant k_{SS} is obtained from Solc and Stockmayer [19] by reducing k_{Smol} by the active fractional area of the encounter sphere. This area is estimated by considering TRAP to be a thick cylinder (height 25 Å, inner radius 17 Å, outer radius 36 Å) within the 36 Å encounter sphere. These dimensions are estimated from the holo TRAP crystal structure 1QAW [20] with a ~15% adjustment to yield a more bloated donut in the apo form (evidenced by dynamic light scattering, data not shown). The binding sites reside on the top surface of the cylinder with a total area of $A = \pi(36\text{-}17)$ Å^2 = 3162 Å^2. This estimates 3162 Å^2 / $4\pi(36\text{Å})^2$ = 3162 / 16278 = 19.4% of the surface of the encounter sphere is eligible for tryptophan binding. The third and most accurate model accounts for the rotational diffusion of TRAP in solution. This rotation dictates that the active surface of TRAP effectively presents itself to a larger fraction of the encounter sphere than if it were not rotating. The rate constant k_{Shoup} employs the model of Shoup, et al. [18] reducing k_{Smol} by $\text{sqrt}(A/\pi)/(\pi R_{Enc}) = 28\%$.

The diffusion constants for apo TRAP and tryptophan are measured to be $480\text{x}10^{-9}$ cm^2/sec and $5010\text{x}10^{-9}$ cm^2/sec at 298K using dynamic light scattering (Sachleben) and NMR (Kleckner) respectively. The temperature-dependence of both diffusion constants is estimated using the Stokes-Einstein relation and published values of water viscosity at each temperature [21].

$$D(T) = \left(\frac{k_B}{6\pi}\right)\left(\frac{T}{R_H \eta(T)}\right)10^4$$

k_B is the Boltzmann constant $1.23x10^{-23}$ J/K, T is the temperature in Kelvin, η(T) is the water viscosity at temperature T in Nsec/m^2, R_H is the hydration (Stokes) radius in meters and 10^4 yields D in cm^2/sec. The encounter rate estimates are combined with the observed rate of binding to yield the probability of tryptophan binding upon encounter: $Pr(Bind) = k_1^f / k_D$ where k_1^f is the first forward rate constant from stopped-flow and k_D is the calculated encounter-rate constant of Trp and TRAP ($k_D = k_{Smol}$, k_{SS} or k_{Shoup}). Figure 8 plots Pr(Bind) at the seven temperatures at which SF data are available and indicates that once tryptophan encounters the TRAP binding site there is only a 0.005 – 0.015% probability of binding.

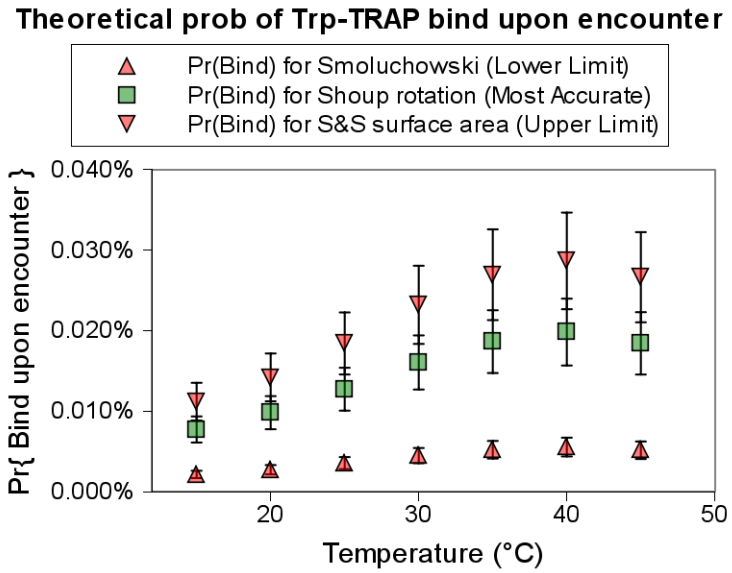


Figure 8. Probability of tryptophan binding upon encounter with TRAP. The upper, lower and most accurate theoretical calculations indicate there is only a 0.005 – 0.03% probability of binding upon encounter (i.e., only 1-in-20000 to 1-in-3333 encounters leads to binding).

This low binding probability may be a result of gating by the TRAP protein and/or the need for waters to exit the binding pocket and/or a rare tryptophan orientation required for insertion. Investigation of these hypotheses will be pursued in future studies.

Tryptophan binding yields step-wise compression of TRAP: Arrhenius analysis of rates

The goal of this section is to thermodynamically describe the five states in M1 (three ground states T+W, TW and TW* and two transition states *[TW]* and *[TW*]*). This is accomplished via Arrhenius and van't Hoff analyses of the temperature-dependence of the SF data [12]. The referenced equations (i.e., using K/K_{Ref} or k/k_{Ref}) are used to directly compare dimensionless quantities representing relative rates and affinities for each step [22]. Temperature-dependence of the four kinetic rates (k_1^F, k_1^B, k_2^F and k_2^B) are interpreted using extended Arrhenius analysis and yields activation enthalpies ($\Delta H^‡$), entropies ($\Delta S^‡$) and heat capacities ($\Delta C_p^‡$) [23, 24]. The significant curvature of the Arrhenius plots (especially step 1F in figure 9) require the use of a non-zero activation heat capacity $\Delta C_p^‡$ term which imparts a linear temperature-dependence on the activation enthalpy $\Delta H^‡$. Statistical considerations for selection of this model are discussed in the appendix.

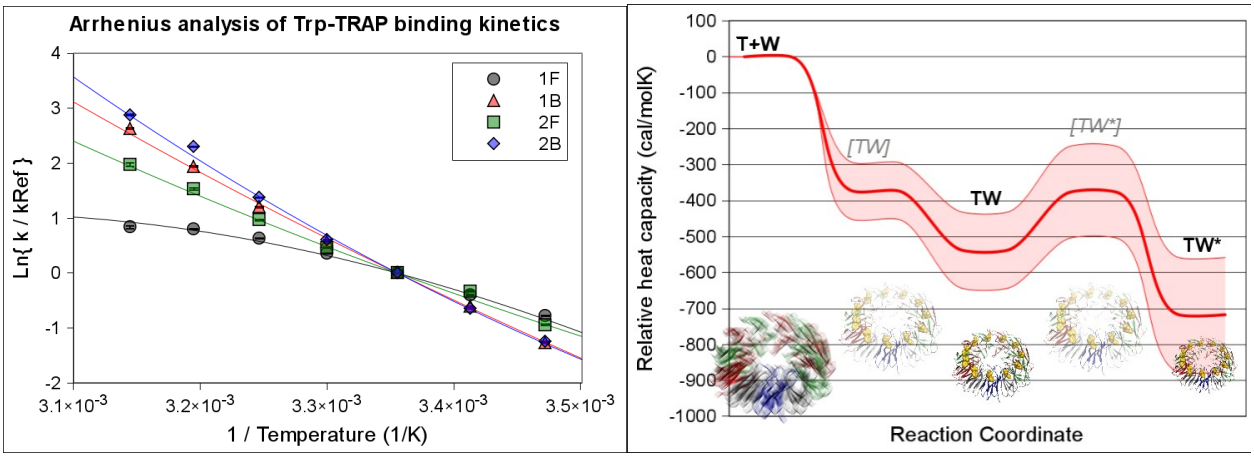


Figure 9. Extended Arrhenius analysis of kinetic data reveal step-wise structural compression of TRAP upon binding tryptophan (reduction in C_p reflects compression of structure). **Left:** Kinetic rate constants for each step are fit to extended Arrhenius equation and exhibit significant curvature (i.e., non-zero $\Delta C_p^‡$). The fitted parameters are $\Delta H^‡$ and $\Delta C_p^‡$. **Right:** Relative C_p with error bar along reaction coordinate informs on changes in structure and dynamics between the five indicated states (the two transition states are bracketed with gray italics). The reduction in C_p is interpreted as a structural compression of TRAP. This is effect is exaggerated in an illustration

at the bottom of the figure for each of the five states. Surprisingly, there is an increase in C_p upon transition between the two bound states TW and TW*.

Extended Arrhenius with sign of $\Delta C_p^{\ddagger}$ inverted from Goldberg & Baldwin (1998)

$$\ln\left(\frac{k}{k_0}\right) = \frac{\Delta H_0^{\ddagger}}{RT_0} - \left(\frac{1}{RT}\right)\left(\Delta H_0^{\ddagger} + \Delta C_p^{\ddagger}\left(T - T_0 - T\ln\left(\frac{T}{T_0}\right)\right)\right)$$

with relative activation entropy

$$\Delta S_{0,\mathrm{Rel}}^{\ddagger} = R\ln(k_0) + R\ln\left(\frac{h}{k_B}\right) + R\ln\left(\frac{1}{T_0}\right) + \frac{1}{T_0}\Delta H_0^{\ddagger} - R\ln(\kappa)$$

Upper-limit of $\Delta S_0^{\ddagger}$ using $\kappa = 1$

Table 2. Extended Arrhenius analysis fit results (T_{Ref} = 298K)

	1F	1B	2F	2B
$\Delta H_{Ref}^{\ddagger}$ (kcal/mol)	12.4 ± 0.54	22.24 ± 0.57	16.73 ± 0.63	23.64 ± 0.78
$\Delta S_{Rel,Ref}^{\ddagger}$ (cal/mol/K)	9.32 ± 1.81	18.34 ± 1.91	-3.99 ± 2.11	16.19 ± 2.62
$\Delta C_p^{\ddagger}$ (cal/mol/K)	-373.6 ± 79.3	163.8 ± 70.34	160.9 ± 71.33	340.6 ± 94.5

The relative reference activation entropy $\Delta S_{Rel,Ref}^{\ddagger}$ is estimated assuming a value of transmission coefficient $\kappa = 1$. This assumption yields a theoretical upper-limit for activation entropies and therefore implies that the activation free energies ($\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger}$) are lower-limits. The value of $\Delta S_{Rel,Ref}^{\ddagger}$ for each step may not be accurate because if κ is less than one then each entropy will be reduced. However, since all entropies are reduced by the same amount ($R\ln(\kappa)$), the differences in entropies are accurate (but not the ratios of entropies). The value of κ is estimated to be in the range of 0.1 to 1.0 [12] and at this lower estimate, each activation entropy is reduced by $R\ln(\kappa) = R\ln(0.1) = 4.5$ cal/mol/K.

Heat capacity (C_p) quantifies the 'extent of states' sampled by the system and is quantitatively defined via the mean fluctuations in enthalpy or entropy [25].

$$C_p = \frac{\langle \delta H^2 \rangle}{kT^2} = \frac{\langle \delta S^2 \rangle}{k}$$

The 'states' may refer to directly observable isomerizations in a reaction mechanism and/or micro-fluctuations within a defined species. Typically, C_p is simply interpreted as a reporter of solvent-accessibility but contributions from protein, ligand and solvent is not well-defined [25]. The following analysis regards C_p loss as a metric for structural compression of TRAP.

Figure 9 illustrates the free state T+W has the highest heat capacity and is set to zero relative C_p for reference. Tryptophan binding yields step-wise compressions of TRAP structure in toward the ground states TW and TW*. The first transition state [TW] reflects a level of intermediate compression which quickly occurs upon Trp-binding. Surprisingly, conversion between the two bound states involves a higher-C_p intermediate [TW*]. This is interpreted as a structural expansion or partial unfolding between the two bound states. This description is conceptually consistent with the MRD results which indicate that apo TRAP samples many more high-energy microstates than holo TRAP (which is a mixture of TW and TW*) at the μs-ms timescale. It is also consistent with dynamic light scattering measurements of the TRAP diffusion coefficient which predict apo TRAP to be about 15% larger than holo TRAP (data not shown).

Prior studies on the tryptophan + TRAP interaction by our lab [4] utilized the wild-type protein and calculated a ΔC_p of -370 cal/molK via ITC. McElroy also estimated the limits of C_p change using the program STC [26]: the minimum magnitude -88 cal/mol/K is simply burial of tryptophan and the contacting TRAP residues. The maximum magnitude -427 cal/mol/K assumes folding of the loop residues in TRAP. These results are smaller than the -636 cal/mol/K observed here for the effective equilibrium (see below). The reason for this discrepancy could be due to any/all of the following: (1) ITC and fluorescence utilize different physical observables which are each subject to different discrepancies from the analysis model. That is, ITC directly measures heat from ligand, protein and solvent while tryptophan fluorescence probes the local environment of the tryptophan and *infers* the heat via van't Hoff analysis [13, 17]. (2) the STC program estimates C_p changes via changes in solvent accessible surface area and disregards contributions from changes in protein, ligand and solvent entropy (3) the TRAP constructs used: WT in McElroy's study and A26I mutant used here.

Comparisons of WT and A26I TRAP constructs reveal that A26I TRAP binds tighter and slower (data not shown). This is due mostly to the 1.5-2.0x greater binding rate k_1^F of the WT protein. This is consistent with a model in which an isoleucine-bearing BC loop gate retards the rate of opening and closing (via larger moment arm) and is also able to make more energetically favorable interactions upon burial from solvent once the tryptophan enters the binding site, compared to an alanine.

Bound state #2 (TW) is more populated, more bonded and more ordered: van't Hoff analysis*

The goal of this section is to analyze the changes in enthalpy, entropy and Gibbs energy coincident with tryptophan binding. In this Arrhenius analysis above, the $\Delta G^{\ddagger}$ values for forward

or reverse steps have not been estimated because this requires assigning a value of transmission coefficient κ hidden in the Arrhenius pre-factor A. κ is often set to unity but this violates the assumption of reversibility in the reaction mechanism [12]. However, the *ratios* of forward and backward rate constants (or equivalently *differences* in $\Delta G^{\ddagger}_{F}$ and $\Delta G^{\ddagger}_{B}$) eliminates the κ term. This provides a sound theoretical basis for obtaining equilibrium constants between ground-states (n.b., $K_i = k_i^F/k_i^B = \exp(-\Delta G_i/RT)$ and $K_{Eff} = ([TW]+[TW^*]) / ([T][W]) = K_1(1+K_2)$). The van't Hoff plots illustrate a significant curvature (figure 10, left); this is expected considering the Arrhenius results since $\Delta C_p = \Delta C_p^{\ddagger}{}_F - \Delta C_p^{\ddagger}{}_B \neq 0$ for both steps.

Table 3. Equilibrium constants from ratios of raw rate constants

Temp (ºC)	K1 (/M)	K2 (-)	K Eff. Bound (/M)	K_d Eff. (μM)
15	289,587 ± 2,504	6.169 ± 0.079	2,076,185 ± 30,085	0.482 ± 0.007
20	217,457 ± 1,451	6.260 ± 0.062	1,578,831 ± 17,684	0.633 ± 0.007
25	176,829 ± 1,223	4.552 ± 0.047	981,696 ± 11,179	1.000 ± 0.012
30	144,420 ± 968	3.984 ± 0.042	719,799 ± 8,171	1.400 ± 0.016
35	99,725 ± 677	3.026 ± 0.034	401,502 ± 4,604	2.500 ± 0.029
40	56,558 ± 705	2.116 ± 0.035	176,218 ± 3,177	5.700 ± 0.102
45	29,387 ± 722	1.848 ± 0.055	83,679 ± 2,810	12.000 ± 0.401

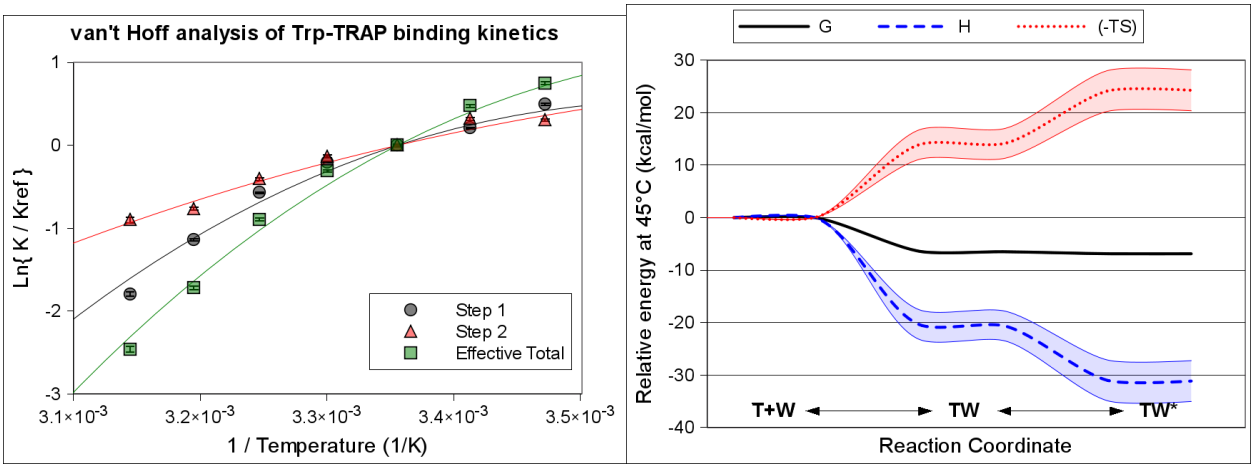


Figure 10. Extended van't Hoff analysis of equilibrium constants provides thermodynamic insight via deconstruction of ΔG into ΔH and TΔS components. **Left:** Equilibrium constants ($K_i = k_i^F/k_i^B$) for both steps and effective total bound $K_{Eff} = K_1(1+K_2)$ are fit to extended van't Hoff

equation and exhibit significant curvature (i.e., non-zero ΔC_p). The fitted parameters are ΔH and ΔC_p. **Right:** Energetic contributions from free energy G, enthalpy H and entropy TS along reaction coordinate. In general, the states proceed to more ordered and more bonded configurations from T+W to TW to TW*. Shaded regions enclose propagated error from fitted parameters in ΔH and ΔC_p the van't Hoff analysis.

Extended van't Hoff from Naghibi, et al. (1995)

$$\ln\left(\frac{K}{K_0}\right) = \frac{\Delta H_0 - T_0 \Delta C_p}{R}\left(\frac{1}{T_0} - \frac{1}{T}\right) + \frac{\Delta C_p}{R}\ln\left(\frac{T}{T_0}\right) \text{ with } K = \frac{k^F}{k^B}$$

Table 4. Extended van't Hoff analysis fit results (T_{Ref} = 298K)

	Step 1	Step 2	Effective Bound
ΔH_{Ref} (kcal/mol)	-9.86 ± 0.97	-7.04 ± 0.93	-15.52 ± 1.18
ΔS_{Ref} (cal/mol/K)	-9.10 ± 3.25	-20.61 ± 3.13	-24.68 ± 3.96
ΔC_p (cal/mol/K)	-535 ± 133	-177 ± 125	-636 ± 159

The results of the extended van't Hoff analysis [13] are shown in figure 10 and table 4. Unlike the raw kinetic rates, this analysis permits accurate calculation of the ground-state ΔG and $T\Delta S$ in addition to ΔH and ΔC_p. Both steps are exothermic ($\Delta H < 0$), enthalpically driven ($\Delta H < T\Delta S$) and have a net decrease in entropy ($T\Delta S < 0$). This reflects a progression of increasing bondedness and increasing order along the reaction coordinate and as temperature increases, both steps are less favored (ΔG is less negative).

*Distinguishing the roles of the two bound states TW and TW**

The first step is clearly tryptophan binding, but the second step remains enigmatic in structural detail. In order to understand what is happening during the second step TW→TW*, it is important to identify the differences between these states. (1) TW* has 177 cal/mol/K less heat capacity than TW. This may reflect a combination of (a) reduced solvent accessible polar surface area in TW*, (b) increased solvent accessible non-polar surface area in TW*, (c) reduced

motions in the protein and ligand in TW* [25]. (2) The TW* state has lower entropy (more ordered) and lower enthalpy (more bonded), both of which are more pronounced at higher temperatures (i.e., $\Delta C_p^{(2)} = -177$ cal/mol/K < 0). Considering the increased order and structure of the TW* state, it seems the most likely candidate for RNA binding (though TW may bind RNA as well). Therefore an upper limit on the activation time is estimated as 0.2-3.0 sec from 15-45°C after binding (k_2^F) even if TW* is the only activated state.

The TW* state is energetically favored over the TW state by a modest 1.8-6.2x from 45-15ºC. Thus, the TW state represents a significant fraction of the bound-state population: 15-35% at 15-45°C and an estimated 45% at 60°C which is the native temperature for *B. stearothermophilus*. This indicates the importance of identifying the roles of each state and what this exchange may imply. The presence of two bound states confers a 3- to 7-fold increase ($1+K_2$) in the effective tryptophan binding equilibrium constant in the range of 45-15ºC. The effective equilibrium compares total bound versus total free $K_{Eff} = ([TW]+[TW^*]) / ([T][W]) = K_1(1+K_2)$; see table 3. The kinetic data indicate a TW↔TW* exchange rate $k_{ex} = k_2^F + k_2^B \sim 0.2 – 4.9$ /sec from 15-45°C, perhaps to permit exit of tryptophan via TW*→TW→T+W. Once in state TW the system can approach one of two transition states, both of which incur an equivalent increase in heat capacity of $\Delta C_p^{\ddagger}{}_{2F} \approx \Delta C_p^{\ddagger}{}_{1B} \approx 160\pm70$ cal/mol/K which is interpreted as a small structural expansion. One transition state [TW] is on-pathway to tryptophan release (TW→T+W) and is less bonded ($\Delta H^{\ddagger}$ more by 5.5 kcal/mol) and less ordered ($\Delta S^{\ddagger}$ more by 22.2 cal/mol/K) than the alternative transition state [TW*] which leads to the putatively active TW* state. Perhaps a very dynamic apo TRAP cannot form a very rigid holo TRAP, but instead the holo TRAP exhibits some (albeit slower) dynamic motions which are required to permit tryptophan release. Furthermore, if these motions involve the aromatic histidine side chains in the loop near

the tryptophan binding site the fluorescence of the bound state may be altered, thus explaining the TW ↔ TW* exchange on the sec timescale. This is also consistent with the observation of two Ile26δ1 signals from holo TRAP in the NMR spectra (figure 1, "*"). Additional analyses indicate that the Ile26δ1 exchange does not directly correspond to the TW ↔ TW* exchange but they may be related qualitatively (data not shown). These observations support a model in which the apo protein is flexible on the μs-ms timescale to prevent RNA from binding and the holo protein is at least locally flexible at the sec timescale to permit Trp release.

Conclusion

This study details some mechanisms by which a protein may modulate its activity to regulate genetic information. The function of apo TRAP is (1) to be able to bind tryptophan and (2) to be *unable* to bind the target RNA (figure S1). The function of holo TRAP is (1) to be able to bind RNA and (2) to be able to release bound tryptophan. The results discussed here indicate TRAP performs these functions via intrinsic ~μs flexibility of the apo state, Trp-induced changes in local structure and concomitant ~μs-ms rigidity with novel ~sec flexibility in the holo state.

This is the first report to quantify site-specific TRAP dynamics at the μs-ms timescale. NMR RD experiments indicate the apo protein undergoes local structural fluctuations in the core and surface at 1-1.5 kHz and 2.5 kHz respectively but the base appears to be rigid. Upon tryptophan binding the local structure is altered in a regulatory "hot spot" and the protein becomes rigid about this activated structure. This supports the putative linkage of dynamics and function in this protein.

The is also the first report on the kinetics of tryptophan binding to TRAP. Stopped-flow fluorescence experiments reveal at least two kinetic steps interpreted as (1) tryptophan binding,

followed by (2) TRAP isomerization. Analysis of the first step indicates tryptophan encounters the TRAP binding site 10^4 times before it is permitted to enter. The second step reports a relatively slow equilibrium between two distinct Trp-TRAP states "TW" and "TW*" ($k_{ex} = k_2^F + k_2^B \sim 0.2 – 4.9$ /sec at 15-45°C). The second bound state TW* is more bonded more ordered, more populated than TW. There is also a step-wise compression of the TRAP structure along the reaction coordinate from free TRAP→TW→TW* and inter-conversion of TW and TW* involves a structurally-bloated intermediate. TRAP is activated in less than 0.2-3.0 sec from 15-45C after binding since k_2^F sets an upper limit assuming that TW* is the only RNA-active conformer.

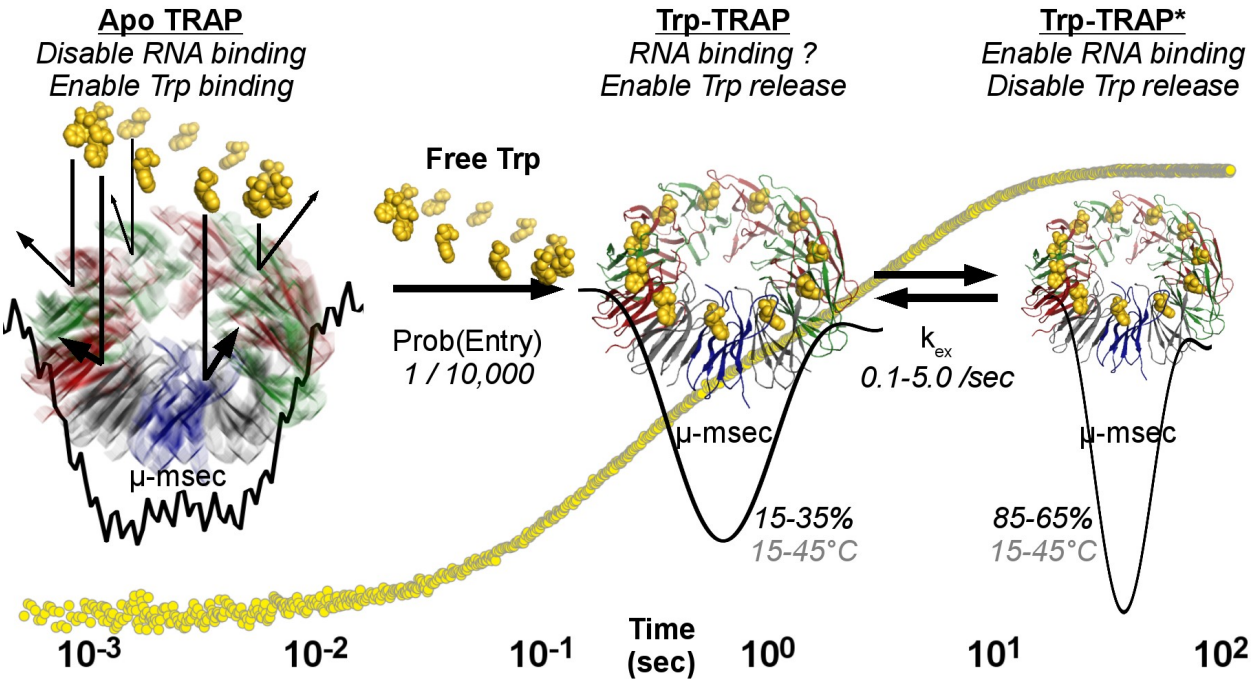


Figure 10. Tryptophan binding alters the conformational landscape of TRAP in fulfillment of biological function. (1) Apo TRAP is bloated and flexible at the μs timescale and Trp encounters the binding site 10^4 times before it is permitted to enter the protein. (2) Upon entry, the structure of TRAP is compressed and initial Trp-TRAP bonds are formed. (3) A new equilibrium is achieved between two Trp-TRAP states such that the second bound state Trp-TRAP* is more populated, more compressed, more bonded, more ordered and rigid at the μs-ms timescale. (4) Exchange between these two bound states Trp-TRAP ↔ Trp-TRAP* reflects intrinsic flexibility

of the Trp-binding loop of the holo protein at the sec timescale which putatively permits bound Trp to exit the protein.

These novel results afford a wealth of information towards understanding the complex relationship between TRAP dynamics and the rate of tryptophan binding and subsequent activation. This is an exciting step towards an increasingly accurate and complex model for this elegant system. Future studies will probe the limits of the putative linkage between structure, dynamics and function by employing the TRAP mutant T28V which is able to bind RNA in the absence of tryptophan [27]. The model proposed here predicts that T28V TRAP is rigid at the μs-ms timescale even without tryptophan since it is already prepared for RNA-binding.

Acknowledgments

Thanks to the Ohio State University, the Council of Graduate Students, the Graduate School, and the Office of Research for sponsoring the Hayes Graduate Research Forum and providing an excellent opportunity for professional development in the pursuit of science. Thanks to Dr. Paul Gollnick and Mr. Brandon Stilb (SUNY at Buffalo) for sample preparation. Thanks to Dr. Lewis Kay and Dr. Philipp Neudecker (Univ. of Toronto), Dr. Chunhua Yuan (OSU CCIC) and Dr. Joseph Sachleben (OSU) for aid in NMR experiments and assignments. Thanks to Dr. Michael Ibba and Ms. Samhita Yadavall (OSU Microbiology) for use of the stopped-flow instrument. Special thanks to Dr. Mark Foster for providing thoughtful input and an enthusiastic working environment. Thanks to the OSU Biophysics program and NIH grant R01GM077234 for funding this project.

References

[1] ,K.Henzler-Wildman ,DKern. Dynamic personalities of proteins. Nature 450 (2007) 964-972.
[2] ,C.Yanofsky. Rna-based regulation of genes of tryptophan synthesis and degradation, in bacteria. RNA 13 (2007) 1141-1154.
[3] ,C.McElroy,A.Manfredo,A.Wendt,P.Gollnick ,MFoster. Trosy-nmr studies of the 91kda trap protein reveal allosteric control of a gene regulatory protein by ligand-altered flexibility. J Mol Biol 323 (2002) 463-473.
[4] ,C.A.McElroy,A.Manfredo,P.Gollnick ,M PFoster. Thermodynamics of tryptophan-mediated activation of the trp rna-binding attenuation protein. Biochemistry 45 (2006) 7844-7853.
[5] ,A.A.Antson,E.J.Dodson,G.Dodson,R.B.Greaves,X.Chen ,PGollnick. Structure of the trp rna-binding attenuation protein, trap, bound to rna. Nature 401 (1999) 235-242.
[6] ,D.S.Wishart ,D ACase. Use of chemical shifts in macromolecular structure determination. Methods Enzymol 338 (2001) 3-34.
[7] ,S.Moon ,D ACase. A new model for chemical shifts of amide hydrogens in proteins. J Biomol NMR 38 (2007) 139-150.
[8] ,P.Lundström,P.Vallurupalli,T.L.Religa,F.W.Dahlquist ,L EKay. A single-quantum methyl 13c-relaxation dispersion experiment with improved sensitivity. J Biomol NMR 38 (2007) 79-88.
[9] ,E.L.Kovrigin,J.G.Kempf,M.J.Grey ,J PLoria. Faithful estimation of dynamics parameters from cpmg relaxation dispersion measurements. J Magn Reson 180 (2006) 93-104.
[10] ,D.M.Korzhnev,P.Neudecker,A.Mittermaier,V.Y.Orekhov ,L EKay. Multiple-site exchange in proteins studied with a suite of six nmr relaxation dispersion experiments: an application to the folding of a fyn sh3 domain mutant. J Am Chem Soc 127 (2005) 15602-15611.
[11] ,P.Kuzmic. Program dynafit for the analysis of enzyme kinetic data: application to hiv proteinase. Anal Biochem 237 (1996) 260-273.
[12] ,D.J.Winzor ,C MJackson. Interpretation of the temperature dependence of equilibrium and rate constants. J Mol Recognit 19 (2006) 389-407.
[13] ,H.Naghibi,A.Tamura ,J MSturtevant. Significant discrepancies between van't hoff and calorimetric enthalpies. Proc Natl Acad Sci U S A 92 (1995) 5597-5599.
[14] ,Y.Liu ,J MSturtevant. The observed change in heat capacity accompanying the thermal unfolding of proteins depends on the composition of the solution and on the method employed to change the temperature of unfolding. Biochemistry 35 (1996) 3059-3062.
[15] ,J.B.Chaires. Possible origin of differences between van't hoff and calorimetric enthalpy estimates. Biophys Chem 64 (1997) 15-23.
[16] ,L.S.Mizoue ,JTellinghuisen. Calorimetric vs. van't hoff binding enthalpies from isothermal titration calorimetry: ba2+-crown ether complexation. Biophys Chem 110 (2004) 15-24.
[17] ,J.R.Horn,D.Russell,E.A.Lewis ,K PMurphy. Van't hoff and calorimetric enthalpies from isothermal titration calorimetry: are there significant discrepancies?. Biochemistry 40 (2001) 1774-1778.
[18] ,D.Shoup,G.Lipari ,ASzabo. Diffusion-controlled bimolecular reaction rates. the effect of rotational diffusion and orientation constraints. Biophys J 36 (1981) 697-714.

[19] ,K.Solc ,W HStockmayer. Kinetics of diffusion-controlled reaction bewteen chemically asymmetric molecules. ii. approximate steady-state solution. Int. J. Chem. Kinet. 5 (1973) 733-752.
[20] ,X.P.Chen,A.A.Antson,M.Yang,P.Li,C.Baumann,E.J.Dodson,G.G.Dodson ,PGollnick. Regulatory features of the trp operon and the crystal structure of the trp rna-binding attenuation protein from bacillus stearothermophilus. J Mol Biol 289 (1999) 1003-1016.
[21] . Crc handbook of chemistry and physics, 90th ed.. David R. Lide (Ed.). Internet Version, 2010.
[22] ,T.Keleti. Errors in the evaluation of arrhenius and van't hoff plots. Biochem J 209 (1983) 277-280.
[23] ,J.M.Goldberg ,R LBaldwin. Kinetic mechanism of a partial folding reaction. 2. nature of the transition state. Biochemistry 37 (1998) 2556-2563.
[24] ,J.M.Goldberg ,R LBaldwin. Kinetic mechanism of a partial folding reaction. 1. properties of the reaction and effects of denaturants. Biochemistry 37 (1998) 2546-2555.
[25] ,N.V.Prabhu ,K ASharp. Heat capacity in proteins. Annu Rev Phys Chem 56 (2005) 521-548.
[26] ,P.Lavigne,J.R.Bagu,R.Boyko,L.Willard,C.F.Holmes ,B DSykes. Structure-based thermodynamic analysis of the dissociation of protein phosphatase-1 catalytic subunit and microcystin-lr docked complexes. Protein Sci 9 (2000) 252-264.
[27] ,P.T.X.Li ,PGollnick. Characterization of a trp rna-binding attenuation protein (trap) mutant with tryptophan independent rna binding activity. J Mol Biol 335 (2004) 707-722.
[28] ,G.Box,J.Hunter ,WHunter. Statistics for experimenters: design, innovation, and discovery , 2nd edition. . John Wiley & Sons, 2005.

Supplementary Information

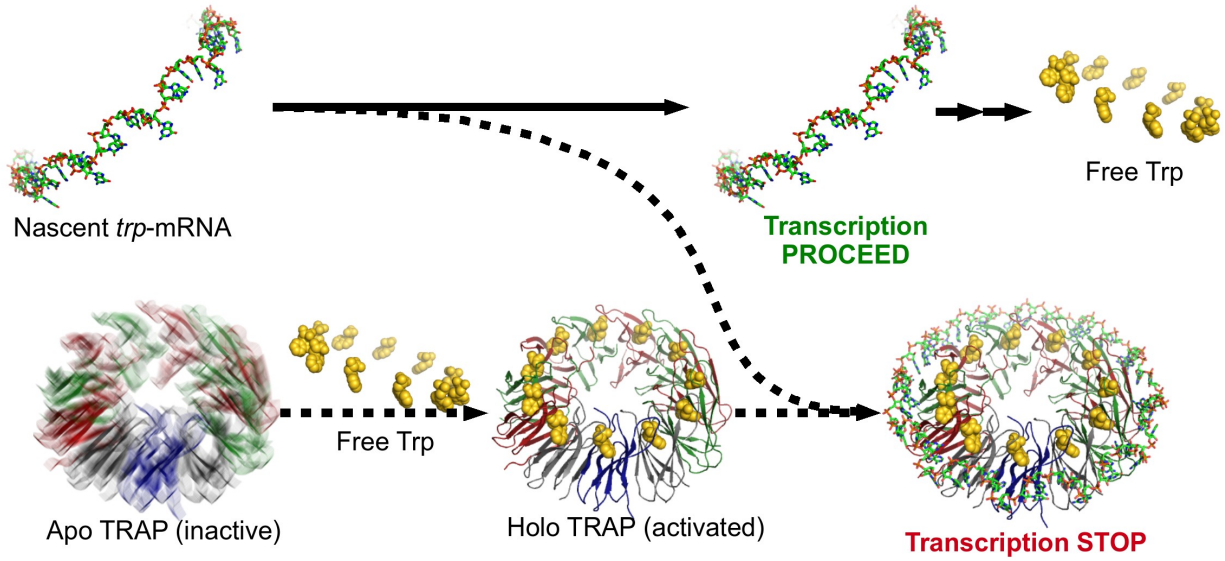


Figure S1. TRAP stops transcription of the nascent trp-mRNA transcript only when activated for RNA-binding by free trytophan. The mechanism of Trp-induced TRAP activation is hypothetically due to a reduction in TRAP's µs-ms dynamics. The 11 identical subunits of TRAP are shown in alternating colors including blue, gray, red and green.

Kinetic model selection for stopped flow data

The goal of this section is to identify the most appropriate mechanistic model for analysis of the Trp-binding kinetic data. The program Dynafit [11] was used since it permits numerical integration of rate laws (i.e., an analytical solution is not required) from an arbitrary reaction mechanism and therefore made this complex model selection tractable. The hypothetical binding mechanisms indicated A through J encompass all 10 possible two- and three-step mechanisms involving either tryptophan binding (B) or TRAP isomerization (I) at each step. For example, model A is "Bind, Bind (BB)" which describes two successive Trp-binding steps with different binding rates.

Kinetic binding mechanisms

$TRAP \overset{Trp}{\Leftrightarrow} TrpTRAP \overset{Trp}{\Leftrightarrow} TrpTrpTRAP$	A [BB]
$TRAP \overset{Trp}{\Leftrightarrow} TrpTRAP \Leftrightarrow TrpTRAP^{*}$	B [BI]
$TRAP \Leftrightarrow TRAP^{*} \overset{Trp}{\Leftrightarrow} TrpTRAP^{*}$	C [IB]
$TRAP \overset{Trp}{\Leftrightarrow} TrpTRAP \overset{Trp}{\Leftrightarrow} TrpTrpTRAP \overset{Trp}{\Leftrightarrow} TrpTrpTrpTRAP$	D [BBB]
$TRAP \overset{Trp}{\Leftrightarrow} TrpTRAP \Leftrightarrow TrpTRAP^{*} \overset{Trp}{\Leftrightarrow} TrpTrpTRAP^{*}$	E [BIB]
$TRAP \overset{Trp}{\Leftrightarrow} TrpTRAP \overset{Trp}{\Leftrightarrow} TrpTrpTRAP \Leftrightarrow TrpTrpTRAP^{*}$	F [BBI]
$TRAP \overset{Trp}{\Leftrightarrow} TrpTRAP \Leftrightarrow TrpTRAP^{*} \Leftrightarrow TrpTRAP^{**}$	G [BII]
$TRAP \Leftrightarrow TRAP^{*} \overset{Trp}{\Leftrightarrow} TrpTRAP^{*} \overset{Trp}{\Leftrightarrow} TrpTrpTRAP^{*}$	H [IBB]
$TRAP \Leftrightarrow TRAP^{*} \Leftrightarrow TRAP^{**} \overset{Trp}{\Leftrightarrow} TrpTRAP^{**}$	I [IIB]
$TRAP \Leftrightarrow TRAP^{*} \overset{Trp}{\Leftrightarrow} TrpTRAP^{*} \Leftrightarrow TrpTRAP^{**}$	J [IBI]

The two models which utilized only isomerization (II and III) were eliminated since they did not invoke a Trp-binding step. Figure S2 illustrates the goodness of fit of each of these models to the binding data at each temperature as well as the sum of all temperatures. The goodness of fit metric is the sum of squares of the error (SSE) which sums the squares of the differences between the observation and the fit (i.e., the residual) at each data point. A smaller values of SSE indicates a more accurate fit to the data and perhaps a more reasonable physical interpretation. Therefore, the goal is to identify the *simplest* model possible which accurately describes the data (i.e., a simple model with small SSE).

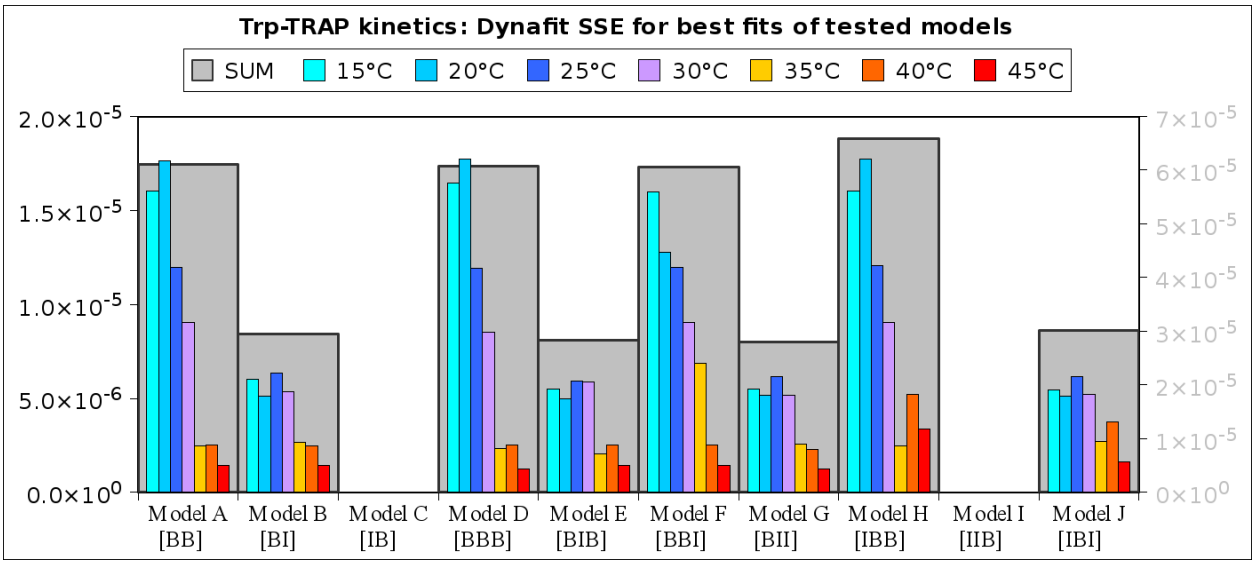


Figure S2. Model selection for interpretation of Trp-binding kinetic data. Model B was selected as most appropriate due to its low value of SSE while invoking the relatively simple two-step model. Thin colored bars represent fits at single temperature and the wide gray bar represents the sum of the fits across all temperatures. SSE values from models C and I are very large and hence are not shown.

Models C and I fit so poorly that their SSE results are not shown in the figure. Models A, D, F and H have the next highest SSE (gray bars) and are therefore removed from further consideration. Models B, E, G and J all have similarly small SSE but since model B is much simpler than the others (i.e., it is two-step instead of three-step) it is selected as the most appropriate framework for further analysis.

Thermodynamic equations for analyses of kinetic data

Thermodynamic relations for both ground and transition‡ energies

$$\Delta G(T) = \Delta H - T\Delta S$$
$$\Delta H(T) = \Delta H_0 + \Delta C_p (T - T_0)$$
$$\Delta S(T) = \Delta S_0 + \Delta C_p \ln\left(\frac{T}{T_0}\right)$$

Heat capacity

$$C_p = \frac{\langle \delta H^2 \rangle}{k T^2} = \frac{\langle \delta S^2 \rangle}{k}$$

Ground-state and transition-state energies

$$\Delta G = \Delta G_F^{\ddagger} - \Delta G_B^{\ddagger} \Rightarrow K = \frac{k^F}{k^B}$$
$$\Delta H = \Delta H_F^{\ddagger} - \Delta H_B^{\ddagger}$$
$$\Delta C_p = \Delta C_{pF}^{\ddagger} - \Delta C_{pB}^{\ddagger}$$

Extended Arrhenius with sign of $\Delta C_p^{\ddagger}$ inverted from Goldberg & Baldwin (1998)

$$\ln\left(\frac{k}{k_0}\right) = \frac{\Delta H_0^{\ddagger}}{R T_0} - \left(\frac{1}{R T}\right)\left(\Delta H_0^{\ddagger} + \Delta C_p^{\ddagger}\left(T - T_0 - T\ln\left(\frac{T}{T_0}\right)\right)\right)$$

Relative activation entropy

$$\Delta S_{0,\mathrm{Rel}}^{\ddagger} = R\ln(k_0) + R\ln\left(\frac{h}{k_B}\right) + R\ln\left(\frac{1}{T_0}\right) + \frac{1}{T_0}\Delta H_0^{\ddagger} - R\ln(\kappa)$$

Upper-limit of $\Delta S_0^{\ddagger}$ using $\kappa = 1$

Extended van't Hoff from Naghibi, et al. (1995)

$$\ln\left(\frac{K}{K_0}\right) = \frac{\Delta H_0 - T_0 \Delta C_p}{R}\left(\frac{1}{T_0} - \frac{1}{T}\right) + \frac{\Delta C_p}{R}\ln\left(\frac{T}{T_0}\right) \text{ with } K = \frac{k^F}{k^B}$$

Equations for model selection

$$\chi^2 = \sum_i \left\{ \left(\frac{Y_i^{\text{Obs}} - Y_i^{\text{Fit}}}{\sigma_i^{\text{Obs}}} \right)^2 \right\}$$

$$\text{Degrees of Freedom (DoF, } \nu\text{)}$$
$$\text{DoF} = \nu = N_{\text{Obs}} - N_{\text{Params}}$$

$$\chi^2_{\text{Red}} = \chi^2 / \nu$$

$$F_0 = \frac{\chi^2_{\text{Red}}(\text{Simple Model})}{\chi^2_{\text{Red}}(\text{Complex Model})}$$

$p - value = Pr(F > F_0)$ Probability simpler model explains the result

$$\text{F is the } \chi^2 \text{ distribution with DoF(Simple) numerator}$$
$$\text{and DoF(Complex) denomenator degrees of freedom}$$

χ^2 is a measure of the goodness of fit. A lower value indicates a better fit of the model to the data. Degrees of freedom indicate the number of independent observations can be used to constrain the model. That is, it requires at least N_{Params} observations to constrain each parameter in the model. After that point, the remaining observations are independent assessments of the fit of the model to the data; the number of these independent assessments is the degrees of freedom.

The χ^2_{Red} is used to normalize the χ^2 goodness of fit to the relative complexity of the model. This affords direct comparison of goodness of fit metrics for models of different complexities. In general, more complex models contain more parameters (N_{Params}), have less degrees of freedom (ν) and fit the data at least as well as the simpler models (smaller χ^2). By comparing χ^2_{Red} values for different models, one can determine if the reduction in degrees of freedom is justified by the subsequent improvement in fit that the more complex model affords.

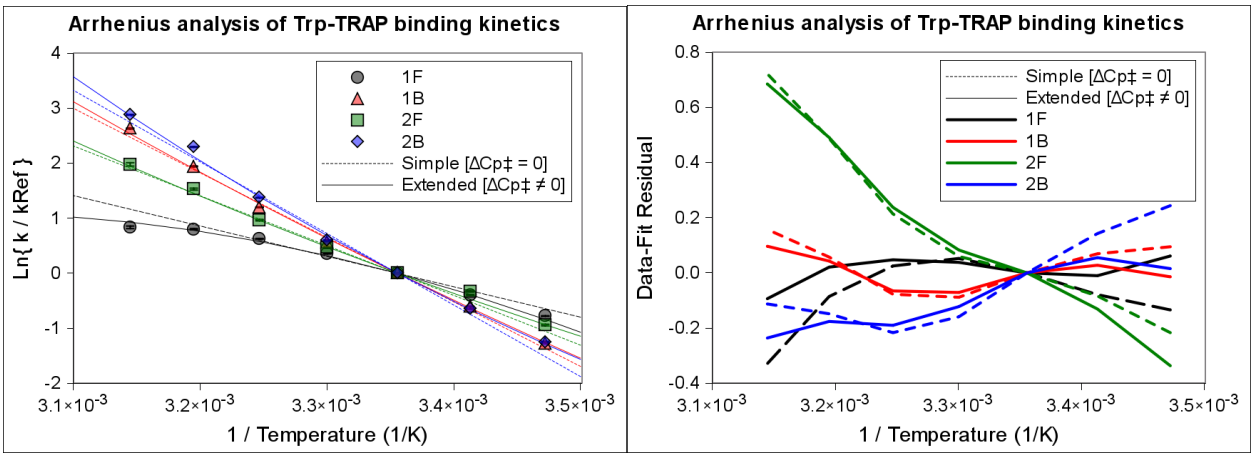
In comparing two models, the ratio of the χ^2_{Red} values (Simple / Complex) is called the F-statistic. A larger value indicates a more strongly that the simpler model is false. The p-value is the probability of observing an F-statistic at least this large even if the simpler model is true. As this probability decreases, it supports the assertion that the simpler model is in fact false. Typically the criterion used to falsify the simpler model is $p < \alpha = 5\%$ or $\alpha = 1\%$ where α is the significance level.

When the p-value is larger than the significance level α, then the simpler model cannot be falsified. However, this is not to say the simpler model is true nor that the more complex model is false. These statistical tests are meant as a quantitative guide to aid in appropriate interpretation of experimental data. The most appropriate interpretation cannot be made without the aid of human judgment and knowledge of the system at hand. For more information, please refer to the excellent book by George P. Box: Statistics for Experimenters [28].

van't Hoff analysis of kinetic rates: extended model is more appropriate than simple model

Extended Arrhenius with sign of $\Delta C_p^{\ddagger}$ inverted from Goldberg & Baldwin (1998)

$$\ln\left(\frac{k}{k_0}\right) = \frac{\Delta H_0^{\ddagger}}{RT_0} - \left(\frac{1}{RT}\right)\left(\Delta H_0^{\ddagger} + \Delta C_p^{\ddagger}\left(T - T_0 - T\ln\left(\frac{T}{T_0}\right)\right)\right)$$



Simple and Extended Arrhenius analyses with T_{Ref} = 298K

Simple model (ΔCp‡ = 0): 1 Parameter and 6 DoF

	1F	E{ 1F }	1B	E{ 1B }	2F	E{ 2F }	2B	E{ 2B }
ΔHRef‡ (kcal/mol)	10.96	1.02	23.27	0.48	17.96	0.42	25.81	0.9
ΔCp‡ (kcal/mol/K)	0	-	0	-	0	-	0	-
Chi^2	1207.92		687.92		326.27		2876.54	
Chi^2 / DoF	201.32		114.65		54.38		479.42	

Extended model (ΔCp‡ ≠ 0): 2 Parameters and 5 DoF

	1F	E{ 1F }	1B	E{ 1B }	2F	E{ 2F }	2B	E{ 2B }
ΔHRef‡ (kcal/mol)	12.4	0.54	22.24	0.57	16.73	0.63	23.64	0.78
ΔCp‡ (kcal/mol/K)	−0.374	0.08	0.16	0.07	0.16	0.07	0.34	0.09
Chi^2	253.01		286.51		153.53		694.97	
Chi^2 / DoF	50.6		57.3		30.71		138.99	

Model comparison

	1F		1B		2F		2B	
F0	3.98		2		1.77		3.45	
p = Pr(F>F0)	7.56%		23.18%		27.36%		9.77%	

In the case of permitting a non-zero value of the activation heat capacity in the Arrhenius analysis (simple vs. complex), the statistical tests indicate p = 7.6%, 23%, 27% and 9.8% for the steps 1F, 1B, 2F and 2B respectively. Although these p-values are larger than the canonical significance level of 5%, the magnitude of the fitted $\Delta C_p^{\ddagger}$ values are sufficiently far from zero to physically justify their use in describing these data. Moreover the van't Hoff analysis strongly

supports use of a non-zero ΔC_p term (e.g., curved van't Hoff plots) and assigning any of the activation $\Delta C_p^{\ddagger}$ values to zero would be inconsistent with this important analysis.

Furthermore, if more data were acquired at intermediate temperatures, the degrees of freedom for both models would increase and thus the models would become more similar in relative complexity. This would reduce the p-values even if there was no absolute improvement in one fit over the other. To reiterate: the most appropriate interpretation cannot be made without the aid of human judgment and knowledge of the system at hand.

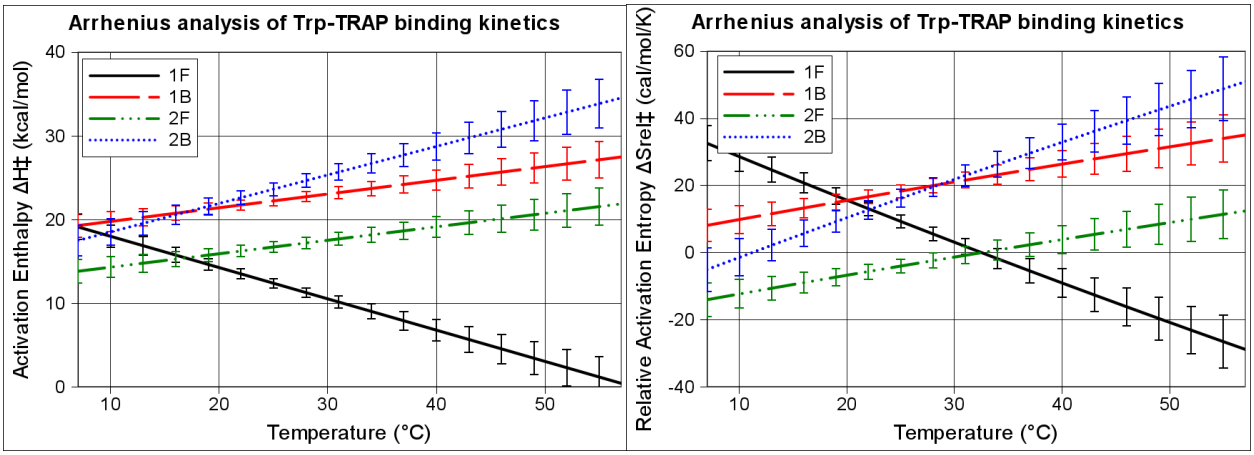


Figure S3. Activation enthalpy and relative activation entropy from Arrhenius analysis. The temperature-dependence of each activation barrier is described by $\Delta Cp^{\ddagger}$. Right: The relative activation entropy $\Delta S_{Rel,Ref}^{\ddagger}$ represents a theoretical upper-limit (κ=1). In practical terms, the scale on the y-axis is accurate but the values may be subject to some offset $R\ln(\kappa)$. Left: The activation enthalpies are not subject to this relative description and are interpreted as absolute values of the enthalpic barriers.

Materials and methods

Sample preparation

TRAP growth and purification were performed in collaboration with Brandon Stilb and Paull Gollnick (SUNY at Buffalo) and described previously [4]. An additional Ile probe was engineered into position 26 (Ala in *B. stearothermophilus* WT) to produce the mutant protein

A26I TRAP. Additional studies report no significant changes in A26I TRAP structure, dynamics and activity compared to WT (data not shown). In this work TRAP refers to *B. stearothermophilus* A26I TRAP unless indicated otherwise. The TRAP sample used for MRD is 1.6 mM monomer and labeled U-[^{2}H, ^{15}N], ILV-[^{13}C^{1}H$_3$] in 500 uL D_2O, 100 mM NaCl, 50 mM $NaPO_4$, 0.02% NaN_3 at pD 8.4 (pH* = 8.0) to minimize the amount of undesired ^{1}H nuclei.

NMR spectroscopy

Assignments of apo and holo TRAP's 29 methyl probes utilized 3D TOCSY correlation of (H^m,N,H^N) and (C^m,N,H^N), 2D and 3D methyl (m) out and back correlation of (H^m,C^m,C^{m-x}), x = {1,2,3} and 2D NOE (C^m, $H^{m\text{-}NOE}$). All NMR spectra were processed and displayed using NMRPipe and NMRViewJ. To generate dispersion curves for each of the 10 MRD conditions , 12 spectra were acquired in random order: ν_{CPMG} = 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1000 Hz (2nd 1000 Hz for error estimation) with CPMG time T = 20 ms and one normalized spectrum with T = ν_{CPMG} = 0. $R_2^{eff}(\nu_{CPMG}) = -\ln\{I(\nu_{CPMG})/I_0\}/T$. I = crosspeak intensity with T and $\nu_{CPMG} > 0$; I_0 = crosspeak intensity with T = ν_{CPMG} = 0.

MRD experiments

Methyl relaxation dispersion curves for each NMR crosspeak are independently analyzed using an in-house developed Matlab program. Intensities of NMR crosspeaks are extracted using the autoFit.tcl package from nmrPipe and errors in intensities are estimated for each dispersion curve using the standard deviation in intensities acquired at repeat ν_{CPMG} values. Effective relaxation rates are calculated as $R_{2,\mathrm{Eff}}^{\mathrm{Obs}}(\nu_{CPMG}) = -\ln\{I(\nu_{CPMG}) / I(T_{CPMG}=0)\}/T_{CPMG}$ where T_{CPMG} is the total CPMG time.

A single dispersion curve is fit to a model of two-state exchange A ↔ B with variable fitting parameters as follows. $|\Delta\omega_H|$ is the magnitude of ^{1}H chemical shift difference between states A and B. $|\Delta\omega_X|$ is the magnitude of ^{A}X chemical shift difference between states A and B (note: in this work, ^{A}X = ^{13}C). P_A is the equilibrium population fraction of state A. $k_{Ex} = k_A + k_B$ is the total rate of exchange between states A and B. k_A and k_B are the rates of exchange from A→B and B→A respectively such with $k_A = (1 - P_A)k_{Ex}$ and $k_B = P_A k_{Ex}$. R_2^0 is the transverse relaxation rate in the absence of exchange assuming $R_2^0{}_A = R_2^0{}_B$.

Provided multiple dispersion curves per NMR crosspeak (different temperatures and/or B_0 fields), the curves are grouped by NMR crosspeak and analyzed as *curve sets*. Each curve set is independently fit to a model of two-state exchange with variable fitting parameters shared as follows. (1) Each crosspeak utilizes a unique $|\Delta\omega_H|$ (fixed to zero for ^{13}C-SQ) and $|\Delta\omega_C|$ expressed in ppm and converted to rad/sec for each dispersion curve via $\Delta\omega_X(\text{rad/sec}) = 2\pi(\gamma_X/\gamma_H)B_0(\text{MHz})\Delta\omega(\text{ppm})$. (2) Each crosspeak's temperature utilizes a unique P_A and k_{Ex} (or equivalently k_A and k_B). (3) Each crosspeak's dispersion utilizes a unique R_2^0. For example, a ^{13}C-SQ data set acquired at 600 MHz and 800 MHz at 25°C, 32°C and 37°C (six dispersion curves in the curve set) are fit by varying $|\Delta\omega_H|$, $|\Delta\omega_C|$, P_A (25°C), k_{Ex} (25°C), P_A (32°C), k_{ex} (32°C), P_A (37°C), k_{ex} (37°C), and six R_2^0 values.

To obtain the most accurate result, the Matlab program employs a grid search over a 4D parameter space encompassing $|\Delta\omega_H|$, $|\Delta\omega_C|$, P_A and k_{Ex}. Each point of this 4D grid is used to create the starting conditions for optimization of the fit using Matlab's *fmincon* function. The same wide grid is used for every curve set: $|\Delta\omega_H|$ ε [0, 1] ppm in 4 steps, $|\Delta\omega_C|$ ε [0.1, 3] ppm in 4 steps, P_A ε [0.5, 0.999] in 4 steps and k_{ex} ε [10, 6000] Hz in 15 steps (a few sets were later re-gridded for more accuracy). P_A and k_{ex} grid points are directly used for the lowest temperature

curves but for subsequent temperatures P_A and k_{Ex} are randomly altered as much as 10% to help direct the fit (as these values likely differ with temperature).

Fit optimization was performed using Matlab's *fmincon* function to minimize the target function:

$$\chi^2 = \sum_{\text{All obs}} \left(\frac{R_{2,\text{Eff}}^{\text{Obs}}(\vec{\text{Obs}}) - R_{2,\text{Eff}}^{\text{Calc}}(\vec{\text{Obs}}, \vec{p})}{\text{Error}\left(R_{2,\text{Eff}}^{\text{Obs}}(\vec{\text{Obs}})\right)} \right)^2$$

$R_{2,\text{Eff}}^{\text{Calc}}$ defined by equations (3.1)-(3.10) in Korzhnev, et al. JACS, 2004. 126(12): p.3964-73.
$\vec{\text{Obs}}$ = Observation conditions: $\nu_{CPMG}, B_{0,}$ Temperature (T)
$\vec{p}$ = Independent fitting parameters: $|\Delta\omega_H|, |\Delta\omega_N|, \{P_A(T)\}, \{k_{\text{Ex}}(T)\}, \{R_2^0(\text{Curve})\}$

Goodness of fit was assessed by visual inspection of the χ^2 map resulting from the grid search. Parameters which are ill-defined are discarded (e.g., fast-exchange limit $k_{ex} >> |\Delta\omega_X|$, poor S/N and/or multiple reasonable fit solutions). However, R_{ex} can almost always be reliably estimated for each curve since dispersion curve fits almost always pass through the observed data: $R_{\text{Ex}} = R_{2,\text{Eff}}^{Fit}(\nu_{CPMG} \approx 0\,\text{Hz}) - R_{2,\text{Eff}}^{Fit}(\nu_{CPMG} \approx 10^4\,\text{Hz})$.

Errors in fitted parameters are estimate by bootstrapping from the residuals as follows. Fit residuals are calculated for each observed ν_{CPMG} value in a given dispersion curve as $\epsilon(\nu_{CPMG}) = R_{2,\text{Eff}}^{\text{Obs}}(\nu_{CPMG}) - R_{2,\text{Eff}}^{\text{Calc}}(\nu_{CPMG})$. The set of residuals is used to create a normal distribution $\boldsymbol{N}(mean(\vec{\epsilon}), var(\vec{\epsilon}))$ for the dispersion curve. A synthetic dispersion *curve* is then created using the best fit at each observed ν_{CPMG} value and a random sample from this normal distribution: $R_{2,\text{Eff}}^{\text{Synth}}(\nu_{CPMG}) = R_{2,\text{Eff}}^{\text{Calc}}(\nu_{CPMG}) + Sample[\,\boldsymbol{N}(mean(\vec{\epsilon}), var(\vec{\epsilon}))\,]$. This is repeated for each curve in the curve set such that a synthetic curve set is produced. The synthetic

curve set is fit using initial conditions which are randomly adjusted as much as 10% from the best fit of the actual data. This process is repeated 100 times to create 100 synthetic curve sets and 100 sets of optimized fit parameters. The error in a given parameter is estimated as the standard deviation of the optimized fit parameter from its 100 element distribution. Errors in subsequent quantities (e.g., k_A, k_B, $\ln(k_A)$, etc.) are estimated using propagation of error assuming all parameters are uncorrelated (zero covariance). Errors in Arrhenius and van't Hoff analyses are estimated from errors from Matlab's fit routine (provided data at more than two temperatures) or propagation of relative error from the fitting variables (when limited to data at only two temperatures).

Provided reasonable fitted exchange rates at multiple temperatures, Arrhenius and van't Hoff analyses are performed. Arrhenius analysis quantifies temperature dependence of exchange *rate* via $\ln(k) = \ln(P) + \left(\frac{-E_A}{R}\right)\left(\frac{1}{T}\right)$ where $k = k_A$ (or k_B) = kinetic rate of exchange from A→B (or B→A), P = Pre-exponential rate (the rate at infinite temperature from A→B (or B→A)), E_A = Activation energy (≈ enthalpy) required to exchange A→B (or B→A), R = gas constant and T = absolute temperature.

van't Hoff analysis quantifies temperature dependence of exchange *populations* via $\ln\left(\frac{k_A}{k_B}\right) = \frac{\Delta S_0}{R} + \left(\frac{-\Delta H_0}{R}\right)\left(\frac{1}{T}\right)$ where $K = k_A / k_B$ = equilibrium constant for exchange, ΔS_0 = system entropy change from state A→B, ΔH_0 = system enthalpy change from state A→B, R = gas constant and T = absolute temperature.