A simple model for determining affinity from irreversible thermal shifts

Justin Hall

Worldwide Medicinal Chemistry, Pfizer, Groton, Connecticut

Correspondence
Justin Hall, Worldwide Medicinal Chemistry, Pfizer, Eastern Point Rd, Groton, CT 06340.
Email: justin.hall@pfizer.com

Abstract
Thermal denaturation ($Tm$) data are easy to obtain; it is a technique that is used by both small labs and large-scale industrial organizations. The link between ligand affinity ($K_D$) and $\Delta Tm$ is understood for reversible denaturation; however, there is a gap in our understanding of how to quantitatively interpret $\Delta Tm$ for the many proteins that irreversibly denature. To better understand the origin, and extent of applicability, of a $K_D$ to $\Delta Tm$ correlate, we define equations relating $K_D$ and $\Delta Tm$ for irreversible protein unfolding, which we test with computational models and experimental data. These results suggest a general relationship exists between $K_D$ and $\Delta Tm$ for irreversible denaturation.

KEYWORDS
activation energy, irreversible denaturation, ligand affinity and unfolding, protein unfolding, thermal denaturation

1 | INTRODUCTION

Thermal melt/thermal shift assays are common in biochemistry; they are experimentally simple, requiring no special protein tags nor immobilization, and the instrumentation is inexpensive. A typical format for this assay is to monitor protein unfolding as a function of temperature; the temperature at which half the protein has unfolded is the thermal melting point ($Tm$). A useful form of this assay is to compare the $Tm$ of a protein between conditions; for example, an increase in $Tm$ after the addition of a putative ligand is a positive indication for ligand binding. The shift between two conditions is called the $\Delta Tm$.

A thermodynamic (reversible) relationship for $\Delta Tm$ as a function of ligand affinity ($K_D$) was described by Brandts and Lin.\(^1\) Unfortunately, temperature-dependent unfolding is not readily reversible for many proteins,\(^2\) which means the Brandt equations cannot be used for many proteins. Thus, instead of a quantitative translation of $\Delta Tm$ to $K_D$, it is instead common for $\Delta Tm$ shifts to be interpreted qualitatively.

There have been several independent publications showing a strong correlate between $\Delta Tm$ and $K_D$.\(^3\)–\(^6\) including publications with hundreds of distinct proteins and ligands.\(^7\)–\(^9\) These data sets contain irreversible unfolding, for which the Brandts equations cannot be applied, yet strangely these irreversible data seem to have the same correlate between $\Delta Tm$ and $K_D$ that is seen for reversible data. So, while it is clear a $\Delta Tm$ to $K_D$ correlate exists for some irreversible denaturation, it is unclear if this is a general feature of irreversible systems.

2 | RESULTS

2.1 | Equations to relate affinity to irreversible thermal denaturation

We will consider denaturation to be irreversible when unfolding is occurring, but the rate of folding is comparatively negligible over the timeframe and temperatures considered.
Example of this could be an unfolded state with a high kinetic barrier for folding, or when aggregation outcompetes folding. To investigate the relationship between \( \Delta T_m \) and \( K_D \) during irreversible denaturation, we considered a protein and ligand system with four simplifying assumptions: (a) changes are two state; (b) the protein can unfold from the apo or the bound state; (c) ligand has a significant binding preference for folded protein; and (d) unfolding is not readily reversible (Figure 1a).

Experimental conditions begin with the protein and ligand in equilibrium, with negligible unfolding or folding, at standard pressure and temperature. Before heating, the free energy change for ligand binding (\( \Delta G_L \)) is:

\[
\Delta G_L = RT \ln \left( 1 + \frac{[L]}{K_D} \right)
\]  

where \( R \) is the gas constant, \( T \) is the temperature in Kelvin, and \([L]\) is the concentration of ligand. The thermodynamic transitions of this system are as follows: apo to unfolded (\( \Delta G_U \)), apo to bound (\( \Delta G_L \)), bound to unfolded-bound (\( \Delta^I G_U \)), and unfolded apo to unfolded-bound (\( \Delta^U G_L \)) (Figure 1a). The thermodynamic cycle is:

\[
\Delta^I G_U - \Delta G_U = \Delta^U G_L - \Delta G_L
\]  

There are two unfolding rate constants for this system, \( k_{U1} \) for apo and \( k_{U2} \) for bound (folding is negligible in this model, thus the folding rate constants are not considered). We assume \( k_{U1} \) and \( k_{U2} \) are temperature-dependent Arrhenius functions:

\[
k = A_0 e^{-E_a/RT}
\]

where \( A_0 \) is a constant relating the approach of the ground state to the transition state and \( E_a \) is the activation energy. \( E_a \) contains both the thermodynamic (\( \Delta G \)) and kinetic energy barrier terms (\( \Delta G^\ddagger \)):

\[
E_a = \Delta G + \Delta G^\ddagger
\]

For our system, the unfolded form of the protein accumulates as a function of \( k_{U1} \) and \( k_{U2} \) multiplied by the concentration of apo or bound protein. For simplicity, we will consider the conditions when the system has either 100% apo or 100% bound (i.e., no ligand, or saturating ligand). Since \( k_{U1} \) and \( k_{U2} \) are Arrhenius functions for the same protein, in the same buffer, at the same heating rates, and over a narrow temperature range relative to absolute, we assume the physical process governing their approach to transition states is the same and therefore their preexponential factors are closely approximate:

\[
A_1 \simeq A_2
\]
constant, $k_{u2}$, at $T_{m,\text{bound}}$ (Figures S1 and S2). Since $\Delta Tm$ is related to these temperatures:

$$\Delta Tm = T_{m,\text{bound}} - T_{m,\text{apo}}$$  \hspace{1cm} (6)

we now set the Arrhenius equations for these two conditions as equal:

$$A_1 e^{-Ea_1/RT_{m,\text{apo}}} = A_2 e^{-Ea_1/RT_{m,\text{apo}} + \Delta Tm}$$  \hspace{1cm} (7)

After simplification of Equation (7) (see Supplemental Materials for detailed derivation), the activations energies can be expressed as:

$$Ea_2 = Ea_1 + Ea_1 \left( \frac{\Delta Tm}{T_{m,\text{apo}}} \right)$$  \hspace{1cm} (8)

The physical difference between $k_{u1}$ and $k_{u2}$, and therefore between $Ea_1$ and $Ea_2$, is the binding of ligand to the protein; thus, Equation (8) says ligand binding increases the activation energy of unfolding by a quantity $Ea_1 \left( \frac{\Delta Tm}{T_{m,\text{apo}}} \right)$ compared to the apo unfolding activation energy. If $\Delta Tm$ is positive, the total activation energy of $Ea_2$ will be greater than $Ea_1$, and $k_{u2}$ will be a smaller rate constant than $k_{u1}$.

For reversible unfolding, positive or negative $\Delta Tm$ is dependent on the binding energy of the ligand for folded or unfolded protein.11,12 To uncover the nature of this dependence for an irreversible system, we return to Equation (2), with the following substitutions:

$$\Delta G_U = Ea_1 - \Delta G^L_U$$  \hspace{1cm} (9)

and

$$\Delta^L G_U = Ea_2 - \Delta^L G^L_U$$  \hspace{1cm} (10)

to find:

$$(Ea_2 - \Delta^L G^L_U) - (Ea_1 - \Delta G^L_U) = \Delta^U G_L - \Delta G_L$$  \hspace{1cm} (11)

If we assume the energy barrier of unfolding for the apo state ($\Delta G^L_U$) and the bound state ($\Delta^L G^L_U$) are closely approximate, then Equation (11) simplifies to:

$$Ea_2 - Ea_1 = \Delta^U G_L - \Delta G_L$$  \hspace{1cm} (12)

Substituting Equation (8) into Equation (12), followed by simplification gives:

$$Ea_1 \left( \frac{\Delta Tm}{T_{m,\text{apo}}} \right) = \Delta^U G_L - \Delta G_L$$  \hspace{1cm} (13)

As with a reversible system, the change in stabilization, $Ea_1 \left( \frac{\Delta Tm}{T_{m,\text{apo}}} \right)$, behaves as the difference of the ligand binding energy for folded and unfolded protein.11,12

Since we have defined the system as having negligible ligand interaction with unfolded protein, Equation (13) simplifies to:

$$Ea_1 \left( \frac{\Delta Tm}{T_{m,\text{apo}}} \right) = - \Delta G_L$$  \hspace{1cm} (14)

The observed $\Delta Tm$ ($\Delta Tm,\text{apo}$) will be the sum of stabilization from ligand binding at the site of interest ($\Delta Tm$), plus any additional extrasite binding ($\Delta Tm,\text{mx}$):

$$\Delta Tm,\text{APP} = \Delta Tm + \Delta Tm,\text{mx}$$  \hspace{1cm} (15)

Substitution from Equations 1, 4, and 15, followed by rearrangement gives:

$$\ln \left( \frac{K_D}{K_D + [L]} \right) = - \frac{Ea_1}{RT} \left( \frac{\Delta Tm,\text{APP}}{T_{m,\text{apo}}} \right) - \ln \left( \frac{K_{DX}}{K_{DX} + [L]} \right)$$  \hspace{1cm} (16)

where $K_{DX}$ is the affinity of extrasite binding. Under conditions of $[L] \gg K_D$, Equation (16) behaves as:

$$\ln \left( \frac{K_D}{[L]} \right) = - \frac{Ea_1}{RT} \left( \frac{\Delta Tm,\text{APP}}{T_{m,\text{apo}}} \right) - \ln \left( \frac{K_{DX}}{K_{DX} + [L]} \right)$$  \hspace{1cm} (17)

which describes the relationship of $K_D$ and $\Delta Tm$ for irreversible denaturation.

In deriving Equation (17), we assumed the ligand binding energy ($\Delta G_L$) will not change with temperature. We also assume there is 100% bound protein when evaluating the expected $\Delta Tm$. We discuss these assumptions below.

Equation (17) is of the $y = mx + b$ form; therefore, a linear relationship is expected for $\ln \left( \frac{K_D}{[L]} \right)$ vs. $\left( \frac{\Delta Tm,\text{APP}}{T_{m,\text{apo}}} \right)$ for irreversible denaturation. To test this, we collected $K_D$ and $\Delta Tm$ for several proteins and ligands with affinities spanning approximately 1,000-fold; these data are linear when plotted as $\ln \left( \frac{K_D}{[L]} \right)$ vs. $\left( \frac{\Delta Tm,\text{APP}}{T_{m,\text{apo}}} \right)$ (Figure 1b). This relationship is consistent with the $K_D$ to $\Delta Tm$ correlate described by others.7-9

We were surprised to discover all the proteins and ligands we tested had similar slopes, ($Ea_1$), and intercepts, $\ln \left( \frac{K_{DX}}{K_{DX} + [L]} \right)$ (Figure 1c). These proteins are not homologs, nor is there anything about them that would lead to the expectation they should have similar $Ea_1$ values (77.9 ± 3.0 kcal/mol); likewise, the ligands are distinct, yet these data also support a similar $\ln \left( \frac{K_{DX}}{K_{DX} + [L]} \right)$ value (−1.9 ± 0.2;
Given the similarity of the slope and intercept for these proteins, and the large amount of irreversible examples previously interpreted by the Brands equations (>100 proteins, >700 ligands for a single protein), these data support a general likeness between proteins for irreversible thermal unfolding and binding. If correct, the relationship expressed in Equation (17), and the empirical fit of \( \frac{E_a}{R} \) and \( \ln \left( \frac{K_{obs}}{K_T + [L]} \right) \), will serve others in prospectively estimating \( K_D \) from \( \Delta T_m \). Consistent with this, we find that affinities calculated from \( \Delta T_m \) agree with affinities calculated by SPR or ITC (Table 1).

2.2 Heating rate and ligand concentration effects

It is important to note that while the observed \( \Delta T_m \) is dependent on ligand concentration and \( K_D \), it is also dependent on the heating rate. For irreversible unfolding, slow heating leads to low \( T_m \) values. To test this effect, we collected experimental data at various heating rates, and calculated the expected \( \Delta T_m \) for these same heating rates through unfolding simulations based on the model depicted in Figure 1a. Our unfolding simulations look as expected for irreversible denaturation, with both the \( T_m \) and the \( \Delta T_m \) decreasing as a function of heating rate, which we also observe experimentally (Figure 2).

In simulations, we observe \( \Delta T_m \) is saturable as the limit of 100% protein occupancy (1:1 stoichiometry); however, when we collect \( \Delta T_m \) while varying ligand concentration, we find increasing the concentration of ligand past the saturation point still increases the \( \Delta T_m \), even for high affinity ligands (ca 0.07 μM) (Figure 3a). The parsimonious interpretation of these data is extrasite binding (>1:1 stoichiometry) occurs as ligand concentration increases. An intriguing alternative explanation of these data is that ligand may be rescuing protein from the apo unfolding path by increasing the effective association rate (e.g., \( k_{on} \times [L] \)) (Figure 1a). To test these hypotheses, we simulated fast and slow ligand association kinetics at variable protein occupancies. Changes in the kinetics of association and dissociation affect the simulated \( \Delta T_m \); however, this effect is also saturable as the limit of 100% protein occupancy and does not explain an increasing \( \Delta T_m \) past saturating ligand concentrations (Figure S3).

To further test ligand effects, we assayed compounds at various concentrations and then fit them to Equation (17).

| Protein | Ligand | \( ^b K_D \) (μM) | \( ^b \)ITC or SPR | \( ^c \)Tm |
|---------|--------|----------------|----------------|--------|
| ^4BASE1 | Cmpd 3 | 14 ± 1.4 | 5.9 ± 0.59 | | |
| Cmpd 4 | 5.4 ± 0.54 | 5 ± 0.5 | | | |
| Cmpd 5 | 4.8 ± 0.48 | 3.1 ± 0.31 | | | |
| Cmpd 6 | 3 ± 0.3 | 1.2 ± 0.12 | | | |
| Cmpd 7 | 2.9 ± 0.29 | 0.9 ± 0.09 | | | |
| Cmpd 8 | 2.9 ± 0.29 | 1.8 ± 0.18 | | | |
| Cmpd 9 | 1.2 ± 0.12 | 2.3 ± 0.23 | | | |
| Cmpd 10 | 0.6 ± 0.06 | 1.4 ± 0.14 | | | |
| ^5cGAS | PF-06928215 | 6.7 ± 0.67 | 3.2 ± 0.32 | | |
| Cmpd 17 | 0.2 ± 0.02 | 0.4 ± 0.04 | | | |
| Cmpd 18 | 2.7 ± 0.27 | 3.1 ± 0.31 | | | |
| Cmpd 19 | 13 ± 1.3 | 12 ± 1.2 | | | |
| ^6NBD1 | ADP | 19 ± 1.9 | 6 ± 0.6 | | |
| AMPNP | 11 ± 1.1 | 5 ± 0.5 | | | |
| ATP | 2.4 ± 0.24 | 1.2 ± 0.12 | | | |
| PhET-ATP | 0.3 ± 0.03 | 0.3 ± 0.03 | | | |

| Cmpd 5 | 4.8 ± 0.48 | 3.1 ± 0.31 | | | |
| Cmpd 4 | 5.4 ± 0.54 | 5 ± 0.5 | | | |
| Cmpd 3 | 14 ± 1.4 | 5.9 ± 0.59 | | | |

^4Standard error is estimated at 10% of values.
^bAffinities were determined by ITC for BASE1 and by SPR for cGAS and NBD1.
^cAffinities were determined without ITC or SPR by SPR for NBD1.

dAfﬁnity was determined by SPR for NBD1.

eWithout cGAS, the global slope and intercept values are −131.2 (±5.3) and −1.9 (±0.20).

We find the slope, \( \left( \frac{E_a}{R} \right) \), and the intercept, \( \ln \left( \frac{K_{obs}}{K_T + [L]} \right) \), change as ligand increases (Figure 3b). Since \( E_a \), the intrinsic unfolding energy of the apo protein, must be independent of ligand concentration, the apparent change in the fit value of \( E_a \) suggests further stabilization is occurring due to extrasite binding. Similarly, without extrasite binding (e.g., \( K_{DX} \gg [L] \)), the intercept should reduce to zero, instead ligand causes a decrease in the intercept (a fit of the intercept for the 50, 100, and 500 μM data sets gives a \( K_{DX} \) of 19, 18, and 24 μM). Therefore, the direction of change in slope and intercept are consistent with this stabilization arising from weak extrasite binding at elevated ligand concentrations. Based on these data, it is not possible to say if this binding occurs at a cryptic site, or at the main site of a partially disrupted protein (either intrinsically or thermally disrupted). Since our model assumes two-state transitions, Equation (17) does not account for intermediate folding effects such as partially disrupted protein populations.
3 | DISCUSSION

We have expressed the expected relationship for \( K_D \) to \( \Delta Tm \) for an irreversible system, which helps explain the strong correlate between \( K_D \) and \( \Delta Tm \) seen in the literature. There is a dissonance in knowingly applying thermodynamic equations to nonequilibrium data; instead of using the Brandt equations, irreversible unfolding data should be interpreted using the \( \ln \left( \frac{K_D}{C} \right) \) vs. \( \frac{\Delta Tm_{app}}{Tm_{app}} \) plots we describe here.

In deriving Equation (17), we have considered the temperature-dependent changes in the activation energies of unfolding of apo (\( E_{a1} \)), bound (\( E_{a2} \)), or ligand binding energy (\( \Delta G_L \)) to be negligible over this temperature range (please note the important difference in maintaining the value of \( \Delta G_L \) at 25°C as constant, not the value of \( K_D \) at 25°C as constant (see Equation (1))). Since we are considering the unfolding of the same protein in apo or bound states, we expect similar temperature dependence to processes where purely protein–protein or protein–solvent interaction are concerned (i.e., all the interactions that do not involve ligand). This logic is manifest in Equation (8), where we find the statement: \( E_{a2} = E_{a1} + E_{a1} \left( \frac{\Delta Tm}{Tm_{app}} \right) \), which says the unfolding process is the same between the apo and bound proteins, but for the additional \( E_{a1} \left( \frac{\Delta Tm}{Tm_{app}} \right) \) term, which from Equation (14), we know to be the ligand binding energy (\( \Delta G_L \)).

We can therefore conclude that, due to the subtraction of bound and apo data, the temperature-dependent changes in the protein unfolding activation energy are not a concern since they are the same for both apo and bound protein. Due to the narrow temperature range considered, it is possible the temperature-dependent change in ligand binding energy is small; however, this is not clear compared to the protein unfolding energy. Therefore, we suspect the temperature-independent assumption for ligand binding energy to contribute the greatest uncertainty to Equation (17).

We have also assumed there is 100% bound protein when evaluating the expected \( \Delta Tm \). If the system does not start with 100% bound, then the experimental \( \Delta Tm \) will be less than the true \( \Delta Tm \) according to the fraction that is apo (see Equation (19) in Methods). Kinetics of ligand association and dissociation may also affect the experimental \( \Delta Tm \), particularly for large energy barriers of association or dissociation (Figure S3). We have attempted to address the occupancy and kinetics effects through unfolding simulations; from these models, we conclude the initial occupancy is the dominant concern compared to ligand binding kinetics. We expect this to be a durable result so long as the fold difference between \( k_{u1} \) and \( k_{u2} \) is small compared to the fold difference between the concentration of apo and bound protein (Figure S3 for further discussion).

To fit thermal denaturation to thermodynamic equations, it is necessary that the unfolding is reversible, and that the heating rate is slow enough that equilibrium is reached between each data point. Any system out of equilibrium cannot be fit by equations that assume complete reversibility has been achieved, and instead should be fit to an incomplete or irreversible model.

The literature suggests the majority of proteins irreversibly denature. Some of these proteins in this study were tested by both \( Tm \) and \( T\text{AGG} \) assays (the latter of which demonstrates irreversible denaturation), and we tested a single protein for irreversible denaturation through heating-rate-dependent changes in \( Tm \) (see Figure 2), but we have not determined if all the proteins included in Figure 1b are irreversibly denatured, yet they still seem to conform to Equation (17). We expect this is due to the intrinsic irreversibility of many of these proteins, but should also result from the rapid heating rate we use. Rapid heating will favor the unfolding rate while limiting folding. In other words, even if a protein is normally reversible, measurements conducted at rates that preclude equilibrium from being reached should cause the system to conform to the framework of the model and derivations presented here.

FIGURE 2  Experimental and simulated irreversible denaturation. Representative data showing the \( Tm \) calculated by simulations using a heating rate of (a) 0.03 or (b) 4°C/min, and experimental data using a heating rate of (c) 0.03 or (d) 4°C/min. Simulated and experimental data are for a \( K_D \) of 0.07 μM at 25°C. Apo (black) and bound (green) protein traces are shown. Protein and ligands are not identified because they are part of an active therapeutic research program at Pfizer. Protein and ligand are the same as used for Figure 3. All data in these figures are available to download.
The similar values seen for proteins in \( \ln \left( \frac{K_D}{T} \right) \) vs \( \Delta \text{Tm} \) plots mean the \( K_D \) to \( \Delta \text{Tm} \) relationship is the same for these proteins. Three reasons occur to us for why there could be a robust \( K_D \) to \( \Delta \text{Tm} \) relationship: (a) most proteins have similar \( Tm_{\text{apo}} \) values on the absolute scale, and \( \Delta \text{Tm} \) experiments compare two conditions over a very narrow temperature range, thus temperature-dependent changes in energies and rates are minimized; (b) \( \Delta \text{Tm} \) is a comparison between the same protein in the same buffer at the same heating rate; this allows an estimation of ligand-dependent effects while minimizing protein-specific information; and (c) there is a fundamental likeness to all proteins; they are polymers of the same 20 amino acids, with generally similar-sized hydrophobic packing spaces separated by hydrophilic secondary structure, this leads to similar protein–solvent interactions when averaged over the whole protein. Points (a) and (c) suggest it is reasonable that \( \Delta \text{Tm} \) data allow for an estimation of the similarities between proteins as they unfold, while points (b) and (c) suggest it is reasonable that \( \Delta \text{Tm} \) data allow for an estimation of the similarities between proteins due to ligand binding.

4 | MATERIALS AND METHODS

4.1 | Experiments

Proteins and ligands are not identified if they are part of an active therapeutic research program. Nucleotide-binding domain 1 (NBD1) of Cystic Fibrosis Transmembrane Conductance Regulator, and the cyclic GMP-AMP Synthase (cGAS) were expressed and purified as described.\(^{15,16}\) Other proteins were purified at Pfizer using standard growth and purification techniques. All the proteins used in these experiments were shown to be greater than 95% pure by gel. Proteins were single domain or multidomain, with molecular weight spans of 20–80 kDa; proteins were either monomers or homodimers.

Ligand affinities for BASE1 are from the work reported by Lo et al.\(^{7}\) SPR affinities for cGAS and NBD1 were determined as described.\(^{15,16}\) Other affinities shown here were determined at 25°C using SPR or estimated from \( K_t \). Activity assays were performed using commercial kits following the manufacturer’s protocol. SPR was performed using a Biacore T200 using CM3 or CM5 chips (GE Healthcare) with immobilized neutravidin to capture enzymatically biotinylated proteins (BirA ligase from Avidity).

Thermal shift assays were done in 20 mM HEPES pH 7.5, 150 mM KCl, and 1 mM TCEP. Assays were performed using a UNiT375 (Unchained Labs) for \( T_{\text{AGG}} \) measurements monitoring static light scattering at 277 nm (SLS 277 nm), or a QuantStudio 6 Flex (Applied Biosystems,
Thermo Fisher Scientific) for Tm measurements monitoring increased fluorescence of SYPRO orange (Sigma). We found the individual values of TAGG and Tm were often several degrees distinct; however, this is typical of these different techniques as they are measures of different phenomena, and the values of ΔTAGG and ΔTm were generally in close agreement. We chose to pursue SYPRO orange Tm data for subsequent experiments since the Tm signal-to-noise was always of a better quality than TAGG.

For our standard experiments, the ligand concentration was 100 μM, heating rate was 4°C/min, and the SYPRO orange concentration was 1:1,000 the manufacturer’s stock (the manufacturer does not provide an absolute concentration, but we used fivefold higher concentration than the recommended 1:5,000 dilution). Ligand concentration was chosen as a practical balance for a strong ΔTm signal against a concern for ligand solubility and extraneous binding. Similarly, our heating rate was chosen to balance a strong ΔTm signal against reductions in signal-to-noise arising from collecting data too rapidly.

Ligands stocks were in DMSO before dilution into protein assay buffer; apo and compound samples were made to contain the same final concentration of DMSO. Variation from our standard collection conditions, such as depicted in Figure 3a, included varying ligand concentration (e.g., between 14 and 448 μM) or heating rate (e.g., between 0.03 and 16°C/min); if not explicitly stated, collection conditions were 100 μM and 4°C/min.

Thermal shift data were imported into GraphPad Prism 7 and analyzed using Equation (18):

\[
y(T) = \left(\frac{y_{\min} - y_{\max}}{1 + e^{\frac{T - T_x}{T_{\text{TAGG}}}}\ln\frac{K_{D\text{X}}}{K_{D\text{X}} + [L]}}\right) + y_{\max} \tag{18}
\]

where \(y(T)\) is the temperature-dependent signal, \(y_{\min}\) and \(y_{\max}\) are fits of the minimal and maximal signal, \(T\) is the temperature, and \(T_x\) is the TTAGG or Tm temperature.\(^\text{18}\) Affinity and ΔTm values were plotted for each protein if the affinity of the ligand was at least fivefold the ligand concentration, the fit of \(\frac{E_{a1}}{RT}\) and \(\ln\left(\frac{K_{D\text{X}}}{K_{D\text{X}} + [L]}\right)\) was determined from a \(\ln\left(\frac{K_{D\text{X}}}{[L]}\right)\) vs \(\frac{\Delta T_{\text{TAGG-max}}}{T_{\text{TAGG-max}}-T_{\text{Tm}}\text{avg}}\) plot. Best-fit analysis was performed in GraphPad Prism 7, values and the standard error are reported. The best fit of \(E_{a1}\) and \(K_{D\text{X}}\) for all proteins and ligands was 77.9 ± 3.0 kcal/mol and 17.9 ± 1.9 μM. All affinity and ΔTm data are available to download.

4.2 | Simulations

In this system, folded protein (F) can be lost to the unfolded form (U), but the unfolded form was treated as effectively irreversible. Folded protein can bind to ligand (FL), and can then unfold (UL), but ligand binding to the unfolded form was unfavorable; thus, the UL complex was treated as immediately decomposing to U. Models were performed where the initial concentration of U was zero, the \([L]\) is much greater than the protein concentration, which was fixed at 10 μM, and generally much greater than \(K_D\).

Initial equilibrium occupancy (\(O_i\)) of FL was determined using Equation (19):

\[
O_i = \left(\frac{1}{\left(\frac{[L]}{K_{D\text{X}}}}{+1}\right}\right) \tag{19}
\]

During simulations, the concentrations of each species can be calculated using these equations:

\[
[U]_t = [U]_{t-1} + ku_1 [F]_t dt + ku_2 [FL]_t dt \tag{20}
\]

\[
[F]_t = [F]_{t-1} + k_{off} [FL]_t dt - ku_1 [F]_t dt - kon[L][F]_t dt \tag{21}
\]

\[
[FL]_t = [FL]_{t-1} + kon[L][F]_t dt - ku_2 [FL]_t dt - k_{off}[FL]_t dt \tag{22}
\]

where species concentration at each time (e.g., [U]₀) is contributed from carryover from the prior time interval (e.g., [U]₀), and an additional contribution from the convertible species (e.g., [F]₀) multiplied by its temperature-dependent rate constant (e.g., \(ku_1\)) over the time interval of integration (dt). Data were calculated using 0.1 s intervals, between 20 and 85°C with variable heating rates.

The value of unfolding rate constants for simulations was estimated based on early experimental data for a protein at 10 μM, with a Tm of ca 50°C, and a heating rate of 4°C/min. Based on this experiment, and a simple two-state unfolding mechanism, we approximated the rate to be around 0.005 s\(^{-1}\) at Tm. This rate gave good visual agreement between experimental and modeled data. Independent of our fit for the model to experimental data, we subsequently found prior work on irreversible unfolding using the same simplified Arrhenius-based model we have used; in their work, these authors fit DSC curves for thermolysin to determine an average rate of 0.004 s\(^{-1}\) at Tm, and an apo activation energy barrier (\(E_a\)) of 67.4 kcal/mol.\(^\text{14}\)

Lacking initial estimates for the preexponential term (\(A_0\)), or \(E_a\), we used 0.005 s\(^{-1}\) to calculate an apparent \(E_a\) of unfolding using the Eyring equation as an approximation for \(A_0\):

\[
k = \frac{k_B T}{h} e^{\frac{-E_a}{RT}} \tag{23}
\]

where \(k\) is the rate constant, \(k_B\) is the Boltzmann’s constant, \(T\) is the temperature in Kelvin, and \(h\) is the Planck’s...
constant.\textsuperscript{10} Using a rate constant approximation of 0.005 s\textsuperscript{-1} at $T_m$, and $k_BT/h$ as the preexponential factor, the activation energy barrier of apo unfolding is ca 22 kca\textsc{l/mo}l.

The $A_0$ and $E_a$ values are lower than we ultimately observed experimentally (ca 1.9 $\times$ 10\textsuperscript{51} s\textsuperscript{-1} and 77.9 $\pm$ 3.0 kca\textsc{l/mo}l), but we chose to use these initial estimates for the remainder of these simulations as we reasoned it should not be necessary to have exact values for comparative analysis between heating rates for apo and bound protein, provided the logic leading to Equation (8) is valid. The values used here, though estimations based on initial data sets, were sufficient for our simulations to predict changes from heating rates or ligand binding. Our simulations, using a 4°C/min heating rate, are available to download.

Using the simulation as described, the heating rate can be changed, and the $T_m$ expected is calculated using numerical integration over time from the initial equilibrium state (Equations (20)–(22)). Using this simulation, we find good agreement between the predicted effect on $T_m$ and $\Delta T_m$ for an irreversible system, and experimental irreversible unfolding (Figure 2). These data demonstrate the protein is undergoing irreversible unfolding,\textsuperscript{13,14} which is consistent with the simple unfolding model these simulations depict (Figure 1a).

ACKNOWLEDGMENTS

The author deeply appreciates the help of Paul Bonin, Reto Horst, Erik Colebrooke Ralph, Michael Schimerlik, and Felix Vajdos during the development of this article.

CONFLICT OF INTEREST

The author declares no conflict of interest.

AUTHOR CONTRIBUTIONS

The work and writing was performed by J.H.

ORCID

Justin Hall \(\text{https://orcid.org/0000-0002-0316-847X}\)

REFERENCES

1. Brandts JF, Lin LN. Study of strong to ultratight protein interactions using differential scanning calorimetry. Biochemistry. 1990;29:6927–6940.

2. Strucksberg KH, Rosenkranz T, Fitter J. Reversible and irreversible unfolding of multi-domain proteins. Biochim Biophys Acta. 2007;1774:1591–1603.

3. Edge V, Allewell NM, Sturtevant JM. High-resolution differential scanning calorimetric analysis of the subunits of \textit{Escherichia coli} aspartate transcarbamoylase. Biochemistry. 1985;24:5899–5906.

4. Hu CQ, Sturtevant JM. Thermodynamic study of yeast phosphoglycerate kinase. Biochemistry. 1987;26:178–182.

5. Manly SP, Matthews KS, Sturtevant JM. Thermal denaturation of the core protein of lac repressor. Biochemistry. 1985;24:3842–3846.

6. Sturtevant JM. Biochemical applications of differential scanning calorimetry. Ann Rev Phys Chem. 1987;38:463–488.

7. Lo MC, Aulabaugh A, Jin G, et al. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. Anal Biochem. 2004;332:153–159.

8. Matulis D, Kranz JK, Salemme FR, Todd MJ. Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor Biochemistry. 2005;44:5258–5266.

9. Pantoliano MW, Petrella EC, Kwasnoski JD, et al. High-density miniaturized thermal shift assays as a general strategy for drug discovery. J Biomol Screen. 2001;6:429–440.

10. Eyring H. The activated complex in chemical reactions. J Chem Phys. 1935;3:107–114.

11. Cimmperman P, Baranauskiene L, Jachimovicute S, et al. A quantitative model of thermal stabilization and destabilization of proteins by ligands. Biophys J. 2008;95:3222–3231.

12. Layton CJ, Hellings HW. Thermodynamic analysis of ligand-induced changes in protein thermal unfolding applied to high-throughput determinations of ligand affinities with extrinsic fluorescent dyes. Biochemistry. 2010;49:10831–10841.

13. Jackson MB, Sturtevant JM. Phase transitions of the purple membranes of \textit{Halobacterium halobium}. Biochemistry. 1978;17:911–915.

14. Sanchez-Ruiz JM, Lopez-Lacomba JL, Cortijo M, Mateo PL. Differential scanning calorimetry of the irreversible thermal denaturation of thermolysin. Biochemistry. 1988;27:1648–1652.

15. Hall J, Brault A, Vincent F, et al. Discovery of PF-06928215 as a high affinity inhibitor of cGAS enabled by a novel fluorescence polarization assay. PLOS One. 2017;12:e0184843.

16. Hall J, Ralph EC, Shanker S, et al. The catalytic mechanism of cyclic GMP-AMP synthase (cGAS) and implications for innate immunity and inhibition. Protein Sci. 2017;26:2367–2380.

17. Hall JD, Wang H, Byrnes LJ, et al. Binding screen for cystic fibrosis transmembrane conductance regulator correctors finds new chemical matter and yields insights into cystic fibrosis therapeutic strategy. Protein Sci. 2016;25:360–373.

18. Senisterra GA, Markin E, Yamazaki K, Hui R, Vedadi M, Awrey DE. Screening for ligands using a generic and high-throughput light-scattering-based assay. J Biomol Screen. 2006;11:940–948.

SUPPORTING INFORMATION

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

How to cite this article: Hall J. A simple model for determining affinity from irreversible thermal shifts. \textit{Protein Science}. 2019;28:1880–1887. \texttt{https://doi.org/10.1002/pro.3701}